

DR WPI; 1999-180971/15  
DR N-PSDB; AAX27234.

XX Delivery of a composition to the central nervous system or spinal  
 PT cord - comprises administration of a non-toxic, proteolytic  
 PT fragment of tetanus toxin in association with a molecule having  
 PT biological function

XX Example 1; Fig 1; 53pp; English.

XX This sequence represents the tetanus toxin fragment C (TTC).  
 CC The invention relates to a method for in vivo delivery of a desired  
 CC composition into a human or animal central nervous system (CNS) or spinal  
 CC cord comprising administering a non-toxic, proteolytic fragment of tetanus  
 CC toxin (TT) in association with at least a molecule having a biological  
 CC function and where the composition is capable of in vivo retrograde  
 CC axonal transport and transsynaptic transport into the CNS or the spinal  
 CC cord of the human or animal and of being delivered to different areas of  
 CC the CNS or the spinal cord. The method can be used for the treatment of  
 CC humans or animals with CNS or spinal cord disease, e.g. neurodegenerative  
 CC and motoneuron diseases such as amyotrophic lateral sclerosis (ALS),  
 CC spinal muscular atrophies (SMA) or neurodegenerative lysosomal storage  
 CC diseases. Compositions comprising hybrid fragments of TT comprising  
 CC fragments C and B can also be used for neuronal mapping and  
 CC immunisations. Use of TT comprising fragments A, B and C results in  
 CC better transport of the fragment inside the organism compared with  
 CC fragment C.

XX Sequence 463 AA;

Query Match Best Local Similarity 100.0%; Score 2456; DB 20; Length 463;

Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSTPIPFSSKNDKCDWNEEDIDVILKSTIINDINNDISDGSFNSVITTPDA 60  
 Db 1 MFSTPIPFSSKNDKCDWNEEDIDVILKSTIINDINNDISDGSFNSVITTPDA 60  
 QY 61 QLVPTNGKAIHLVNNESSEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQYGTN 120  
 Db 61 QLVPTNGKAIHLVNNESSEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQYGTN 120  
 QY 121 EYSIISSMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 180  
 Db 121 EYSIISSMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 180  
 QY 181 FTTTNDRLSSANLYINGVLMGSAETTGATREDNNITLKDRCNNNNQYVSIIDKFRIF 240  
 Db 181 FTTTNDRLSSANLYINGVLMGSAETTGATREDNNITLKDRCNNNNQYVSIIDKFRIF 240  
 QY 241 CKALNPKEIEKLYTSYSTIFLDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLT 300  
 Db 241 CKALNPKEIEKLYTSYSTIFLDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLT 300  
 QY 301 NAPSITNGKLNITYRRLYNGLKFIIRKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYP 360  
 Db 301 NAPSITNGKLNITYRRLYNGLKFIIRKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYP 360  
 QY 361 KQGNAFNNNDRLIRVGYNAAGIPLKKMEAVKLRDLKTYSVQKLKYDDKNNASGLVGTHNG 420  
 Db 361 KQGNAFNNNDRLIRVGYNAAGIPLKKMEAVKLRDLKTYSVQKLKYDDKNNASGLVGTHNG 420  
 QY 421 GOIGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463  
 Db 421 GOIGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463

RESULT 2  
 ID AAF70345 standard; Protein; 573 AA.

XX AAF70345;  
 XX 22-APR-1991 (first entry)  
 XX

DE Portion of B fragment and all of the C fragment of tetanus toxin.

XX TT; vaccine.

XX Clostridium tetani.

XX EP209281-A.

XX 21-JAN-1987.

XX 27-JUN-1986; 86EP-0305029.

XX 28-JUN-1985; 85GB-0016442.

XX (WELL) WELLCOME FOUNDATION LTD.

XX WPI: 1987-015999/03.

XX N-PSDB; AAN70545.

XX Claim 4; Fig 1; 36pp; English.

XX Gene product comprises a tetanus toxin fragment, which may be

XX expressed in a transformed host, and used as an antigen in vaccine

XX production, against the disease.

XX Sequence 573 AA;

Query Match Best Local Similarity 99.8%; Score 2451; DB 8; Length 573;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VFSTPIPFSSKNDKCDWNEEDIDVILKSTIINDINNDISDGSFNSVITTPDAQ 61  
 Db 112 VFSTPIPFSSKNDKCDWNEEDIDVILKSTIINDINNDISDGSFNSVITTPDAQ 171  
 QY 62 LVPNGKAIHLVNNESSEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQYGTNE 121  
 Db 172 LVPNGKAIHLVNNESSEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQYGTNE 231  
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 181  
 Db 232 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWV 291  
 QY 182 ITTNDRLSSANLYINGVLMGSAETTGATREDNNITLKDRCNNNNQYVSIIDKFRIF 241  
 Db 292 ITTNDRLSSANLYINGVLMGSAETTGATREDNNITLKDRCNNNNQYVSIIDKFRIF 351  
 QY 242 KALNPKEIEKLYTSYSTIFLDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 301  
 Db 352 KALNPKEIEKLYTSYSTIFLDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 411  
 QY 302 APSYTGKLNITYRRLYNGLKFIIRKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYP 361  
 Db 412 APSYTGKLNITYRRLYNGLKFIIRKRYTPNNEIDSFVKSQGFITKLYSYNNNEHIVGYP 471  
 QY 362 DQNAFNNNDRLIRVGYNAAGIPLKKMEAVKLRDLKTYSVQKLKYDDKNNASGLVGTHNG 421  
 Db 472 DQNAFNNNDRLIRVGYNAAGIPLKKMEAVKLRDLKTYSVQKLKYDDKNNASGLVGTHNG 531  
 QY 422 OTGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 463  
 Db 532 OTGNDPNRDIILASNMVFNHLKDKIIGCDWYFVPTDEGWTND 573

RESULT 3  
 ID AAB61169 standard; Protein; 1315 AA.

XX

AC AAB61169;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Clostridium tetani TeNT.  
 XX  
 KM Clostridium tetani: TeNT; tetanus toxin; antibacterial; vaccine;  
 XX TeNT fragment C; infection.  
 XX  
 OS Clostridium tetani.  
 PN W0200100839-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000MO-GB02428.  
 XX  
 PR 25-JUN-1999; 99GB-0014861.  
 XX  
 PA (UNLO ) IMPERIAL COLLEGE SCI TECHNOLOGY & MED.  
 XX Fairweather NF, Sinha K;  
 PI WPI; 2001-123014/13.  
 DR  
 XX  
 XX  
 PT New polypeptide, useful for treating infections of Clostridium tetani,  
 PT comprises tetanus toxin fragment with a mutation in a loop region,  
 PS  
 PS Disclosure; Page 39; 43pp; English.  
 CC  
 CC The present sequence is given in a specification relating to a novel  
 CC polypeptide comprising tetanus toxin (TeNT) fragment C or its immunogenic  
 CC fragment, containing a mutation in a loop region. The mutation results in  
 CC a reduction in the binding of TeNT fragment C or its immunogenic fragment  
 CC to gangliosides and primary motoneurons, and/or a reduction in the  
 CC ability of TeNT fragment C or its immunogenic fragment to undergo  
 CC retrograde transport. The polypeptide is useful for treating, preventing  
 CC and reducing the susceptibility to Clostridium tetani infection in a  
 CC human or animal, and also for producing antibodies which recognise groups  
 CC within TeNT polypeptides. Antibody produced against the polypeptide is  
 CC also useful for treating Clostridium tetani infection.  
 XX  
 XX  
 SQ Sequence 1315 AA;

Query Match 99.5%; Score 2444; DB 22; Length 1315;  
 Best Local Similarity 99.8%; Pred. No. 4.2e-189;  
 Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 VESTPIPSYSKNDICWVNEEDIDVILKSTIILNDIINDIISDGSFNSSVITYPDAQ 61  
 DB 854 VESTPIPSYSKNDICWVNEEDIDVILKSTIILNDIINDIISDGSFNSSVITYPDAQ 913  
 QY 62 LVPGINGKAHLVNNESSEVIVHAKMDIEYNDMFNFTVFWLNPVKVSASHLOXYGNE 121  
 DB 914 LVPGINGKAHLVNNESSEVIVHAKMDIEYNDMFNFTVFWLNPVKVSASHLOXYGNE 973  
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKENAVLANKWVF 181  
 DB 974 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKENAVLANKWVF 1033  
 QY 182 ITTNDRLSSANLYINGVLGSAEITGLAIREDNNTTLKDRNNNNNOYVSDIKFRIFC 241  
 DB 1034 ITTNDRLSSANLYINGVLGSAEITGLAIREDNNTTLKDRNNNNNOYVSDIKFRIFC 1093  
 QY 242 KALNPKRIEKLIVSYSTTLRDFWGNPLRYDEYYLIPVASSSKDQVLNITDYMLTN 301  
 DB 1094 KALNPKRIEKLIVSYSTTLRDFWGNPLRYDEYYLIPVASSSKDQVLNITDYMLTN 1153  
 QY 302 APSYTGKLIIVYRRILYNGKFKITKRYTPNNEIDSPVKSDFIKLIVSYNNNEHIVGPR 361  
 DB 1154 APSYTGKLIIVYRRILYNGKFKITKRYTPNNEIDSPVKSDFIKLIVSYNNNEHIVGPR 1213  
 QY 362 DGNAFNNLDRILRVGYNABGIPLYKKMEAVKLRDLKTSYVOLKLYDDKNASGLVGTGNC 421

DB 1214 DGNAFNNLDRILRVGYNABGIPLYKKMEAVKLRDLKTSYVOLKLYDDKNASGLVGTGNC 1273  
 QY 422 QIGNDPRNDIILASNMWYNHLLKDKITLGCDFWVPVPPDEGTND 463  
 DB 1274 QIGNDPRNDIILASNMWYNHLLKDKITLGCDFWVPVPPDEGTND 1315

RESULT 4  
 AAW48909  
 ID AAW48909 standard; Protein: 618 AA.  
 XX  
 AC AAW48909;  
 XX  
 DT 23-SEP-1998 (first entry)  
 XX  
 DE SOD-1/TTC hybrid protein.  
 XX

KM Chimeric: copper-zinc superoxide dismutase; SOD-1; TTC; SOD:TeT451;  
 KM tetanus toxin fragment C; tetanus holotoxin; nerve cell; stroke;  
 KM neurological disorder; oxidative stress; brain hypoxia-reperfusion;  
 XX epilepsy; Parkinson's disease; Huntington's disease.

OS Chimeric - Homo sapiens.  
 XX Chimeric - Clostridium tetani.  
 FH  
 FT Key Location/Qualifiers  
 FT Region 1..163  
 FT /note= "SOD-1"  
 FT Region 168..618  
 FT /note= "TTC moiety"

US5780024-A.  
 14-JUL-1998.  
 XX  
 XX 21-JUN-1996; 96US-0668381.  
 XX  
 XX 23-JUN-1995; 95US-0000473.  
 XX 21-JUN-1996; 96US-0668381.

PA (GENO ) GEN HOSPITAL CORP.  
 PA (UYMA-) UNIV MARYLAND BALTIMORE.  
 XX  
 XX Brown RH, Fishman PS, Francis JW, Hosler BA;  
 DR WPI; 1998-412999/35.  
 XX N-PSDB; AAV32580.

PT New hybrid protein of superoxide dismutase and tetanus toxin  
 PT fragment C - having increased uptake by neurons and retention of  
 PT enzymatic activity in these cells, for treating neurological  
 XX diseases associated with oxidative stress

Claim 7; Columns 23-26; 23pp; English.

CC The present sequence represents an enzymatically active human  
 CC copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl  
 CC terminus with the tetanus toxin fragment C (TTC) moiety. The TTC  
 CC moiety constitutes amino acid residues 865-1315 of the tetanus  
 CC holotoxin. The hybrid protein, referred as SOD:TeT451, is claimed  
 CC to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic  
 CC activity; (b) the TTC moiety selectively binds to nerve cells and  
 CC allows uptake of the hybrid protein into these cells; and (c) it  
 CC retains substantial SOD enzymatic activity following cellular uptake.  
 CC SOD:TeT451 is claimed to be useful for treating neurological disorders  
 CC associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion,  
 CC epilepsy, Parkinson's and Huntington's diseases.

SQ Sequence 618 AA;  
 Query Match 99.3%; Score 2440; DB 19; Length 618;  
 Best Local Similarity 99.6%; Pred. No. 3e-189;

CC The invention relates to a non toxic polypeptide, for delivery of a  
CC therapeutic agent to a neuronal cell, which comprises a binding domain  
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated  
CC as Hc) that binds to the neuronal cell and a translocation domain (amino  
CC terminal half of HC, designated as HN), that translocates the therapeutic  
CC agent into the neuronal cell, where the translocation domain is not a HN  
CC domain of a clostridial neurotoxin and is not a fragment or derivative of  
CC a HN domain of a clostridial toxin. Polypeptides of the invention are  
CC useful for the treatment of a disease state associated with neuronal  
CC cells. The polypeptide constructs are useful for delivering therapeutic  
CC substances to neuronal cells. They are useful to treat disorders of the  
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours  
CC and infection. They are also useful in gene therapy. The present sequence  
CC is modified clostridial heavy chain fragment. This sequence is  
CC constructed by fusing the binding domain of tetanus neurotoxin (TcNT)  
CC with translocation domain of diphtheria neurotoxin.

xx  
SQ Sequence 665 AA:

Query Match 97.6%; Score 2397; DB 22; Length 665;  
Best Local Similarity 100.0%; Pred. No. 1e-185;  
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 12 SKNLDQWVNEEDIVYLKKSTLLINDINNDITSDISGENSSVITYYPDAOLVPGINGKAI 71  
|||||  
Db 212 SKNLDQWVNEEDIVYLKKSTLLINDINNDITSDISGENSSVITYYPDAOLVPGINGKAI 271

QY 72 HLYVNSESSEVYIAHKAMIDIEYNDMFNNFTYSFWLRYVKVSAASHLEQGTNETSYIISSMKKH 131  
|||||  
Db 272 HLYVNSESSEVYIAHKAMIDIEYNDMFNNFTYSFWLRYVKVSAASHLEQGTNETSYIISSMKKH 331

QY 132 SLSTSGWWSYLLKGNLITWLKDSAGEVRQITPTREDLPDKFNAYLANKKWFEITITNDRLLS 191  
|||||  
Db 332 SLSTSGWWSYLLKGNLITWLKDSAGEVRQITPTREDLPDKFNAYLANKKWFEITITNDRLLS 391

CC	The invention relates to a non toxic polypeptide- for delivery of a
CC	therapeutic agent to a neuronal cell, which comprises a binding domain
CC	(carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC	as Hc) that binds to the neuronal cell and a translocation domain (antit-
CC	terminal half of HC, designated as HN), that translocates the therapeut-
CC	agent into the neuronal cell, where the translocation domain is not a Hc
CC	domain of a clostridial neurotoxin and is not a fragment or derivative of
CC	a HN domain of a clostridial toxin. Polypeptides of the invention are
CC	useful for the treatment of a disease state associated with neuronal
CC	cells. The polypeptide constructs are useful for delivering therapeutic
CC	substances to neuronal cells. They are useful to treat disorders of the
CC	CNS including neurodegenerative diseases, stroke, epilepsy, brain tumou
CC	and infection. They are also useful in gene therapy. This sequence is
CC	is modified clostridial heavy chain fragment. This sequence is
CC	constructed by fusing the binding domain of tetanus neurotoxin (TeNT)
CC	with translocation domain of diphtheria neurotoxin.
XX	
SQ	Sequence      665 AA;
	Query Match                  97.6%; Score 2397; DB 22; Length 665;
	Best Local Similarity        100.0%; Pred. No. 1e-185;
	Matches    452; Conservative     0; Mismatches    0; Indels     0; Gaps
OY	12 SKNDCWVNEEDIDYLTKRSTLINDINNDIISDISEGSSVITYPPDAQLVPINGKAI 71
Db	212 SKNDCWVNVED IDYLKKSTLINDDINDIIDSDIGSNESSVITYPDAQLVGINGKAI 271
OY	72 HLAVNESSEVIYHKAMDIEYNDFNNFTVSFWLRPVKSASHLEQGTMEYSIISSMKKH 131
Db	272 HLAVNESSEVIYHKAMDIENDFNFTVSFWLRPVKSASHLEQGTMDYSIISSMKKH 331
OY	132 SLSTSGWSVSLKGNNLIWTLKDSAGEVQIPIPRDLDPKFNAVLANKWFITITNDRLSS 191
Db	332 SLSTSGWSVSLKGNNLIWTLKDSAGEVRQ ITPRDLPKFNAVLANKWFIITTDRLSS 391
OY	192 ANLYINGVILMGSAEITGLAIREDNNTIKLDRCNNNNQYVSIDKRIRICKALNPKEIEK 251
Db	392 ANLYINGVILMGSAEITGLAIREDNNTIKLDRCNNNNQYVSIDKRIFECKALNPKEIEK 451
OY	252 LYTYSYLSTFPLRFQWGNPLKRYDEYYILIPAASSKDQOLNTIDMYLTAPSYTNGKLN 311
Db	452 LYTYSYLSTFPLRFQWGNPLKRYDEYYILIPAASSKDQOLNTIDMYLTAPSYTNGKLN 511
OY	312 IYVRRLYNGLKETIKRTYPPNNEIDSFFKSGDFIKLYSVNNNEHIVGYPRDGNAFNNDLR 371
Db	512 IYVRRLYNGLKETIKRTYPPNNEIDSFFKSGDEFIKLYSVNNNEHIVGYPRDGNAFNNDLR 571
OY	372 ILRVGVNAPGIPLYKKMEAYKLRLDKTYTVQOLKYXDRKNASGLGYGHNOIGDNPNRDI 431
Db	572 ILRVGVNAPGIPLYKKMEAYKLRLDKTYTVQOLKYXDRKNASGLGVTHNGOIGDNPNRDI 631
OY	432 LIASNWFNHLLKDKILGDMWFVPTDGGWTFND 463
Db	632 LIASNWFNHLLKDKILGDMWFVPTDGGWTFND 663
RESULT 6	
AAE07889	
ID	AAE07889 standard; Protein: 882 AA.
XX	
AC	AAE07889;
XX	
DT	01-NOV-2001 (first entry)
XX	
DE	Modified clostridial heavy chain-superoxide dismutase conjugate #1.
KM	Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KM	tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KM	superoxide dismutase; SOD; diphtheria neurotoxin; tetanus neurotoxin;
KM	TENT.
XX	
OS	Chimeric - Bacillus stearothermophilus.
OS	Chimeric - Corynebacterium diphtheriae.



Query Match 97.6%; Score 2397; DB 22; Length 907;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-185;  
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SKNLDGVNDEEDIDVILKSTIILNDIINDIISDISGNSSVITYPDAQLVPGINGKAI 71  
 DB 454 SKNLDGVNDEEDIDVILKSTIILNDIINDIISDISGNSSVITYPDAQLVPGINGKAI 513  
 QY 72 HLNNSESEVIVKAMDIENDMFNFTVSFWLRPKVSAHLEQYGTNEYSIISSMKH 131  
 DB 514 HLNNSESEVIVKAMDIENDMFNFTVSFWLRPKVSAHLEQYGTNEYSIISSMKH 573  
 QY 132 SLISGWSVSLKGNLILWTLKDSAGEVROITPDLDPKFNAYLANKWVFITINDRLSS 191  
 DB 574 SLISGWSVSLKGNLILWTLKDSAGEVROITPDLDPKFNAYLANKWVFITINDRLSS 633  
 QY 192 ANLYINGVLMGSAEITGLAIRREDNNTTLKLDRCNNNNQYVSIDKFRIFCKALNPKIEK 251  
 DB 634 ANLYINGVLMGSAEITGLAIRREDNNTTLKLDRCNNNNQYVSIDKFRIFCKALNPKIEK 693  
 QY 252 LKTSYLSITFLRDFWGNPLKRYDTREYLLIPVASSSKDVOLKNTTDMYLTNAPSTYNGKLN 311  
 DB 694 LKTSYLSITFLRDFWGNPLKRYDTREYLLIPVASSSKDVOLKNTTDMYLTNAPSTYNGKLN 753  
 QY 312 IYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGPKDGNAFNNLDR 371  
 DB 754 IYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGPKDGNAFNNLDR 813  
 QY 372 ILRFGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNRI 431  
 DB 814 ILRFGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNRI 873  
 QY 432 LIASNMYFNHLKDKILGCDWYFVPTDEGWTND 463  
 DB 874 LIASNMYFNHLKDKILGCDWYFVPTDEGWTND 905

RESULT 8  
 AAE07903  
 ID AAE07903 standard; Protein; 1052 AA.  
 AC AAE07903;  
 DT 01-NOV-2001 (first entry)  
 DE C. botulinum C2 translocation domain with TeNT binding domain #2.  
 KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;  
 KM tumour; infection; neurodegenerative disease; gene therapy;  
 KW botulinum neurotoxin; tetanus neurotoxin; TeNT.  
 XX OS Clostridium botulinum.  
 OS Clostridium tetani.  
 XX PN MO200158936-A2.  
 XX PD 16-AUG-2001.  
 XX PF 04-DEC-2000; 2000MO-GB04644.  
 XX PR 02-DEC-1999; 99GB-0028530.  
 XX PR 07-APR-2000; 2000GB-0008658.  
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
 XX PI Shone CC, Sutton JM, Silman N;  
 XX WPI; 2001-514643/56.  
 XX New non toxic polypeptide for delivery of a therapeutic agent for the  
 PT treatment of a CNS disorder comprising a binding domain that  
 PT translocates the therapeutic agent into the neuronal cells -

PS Example 2; Page 50; 50bp; English.  
 XX The invention relates to a non toxic polypeptide, for delivery of a  
 CC therapeutic agent to a neuronal cell, which comprises a binding domain  
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated  
 CC as HC) that binds to the neuronal cell and a translocation domain (amino  
 CC terminal half of HC, designated as HN), that translocates the therapeutic  
 CC agent into the neuronal cell, where the translocation domain is not a HN  
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of  
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are  
 CC useful for the treatment of a disease state associated with neuronal  
 CC cells. The polypeptide constructs are useful for delivering therapeutic  
 CC substances to neuronal cells. They are useful to treat disorders of the  
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours  
 CC and infection. They are also useful in gene therapy. The present sequence  
 CC is C. botulinum C2 enterotoxin translocation domain with tetanus  
 CC neurotoxin (TeNT) binding domain used in the exemplification of the  
 CC invention.  
 CC XX

Sequence 1052 AA;

Query Match 97.6%; Score 2397; DB 22; Length 1052;  
 Best Local Similarity 99.6%; Pred. No. 2e-185;  
 Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 SYKNLDGVNDEEDIDVILKSTIILNDIINDIISDISGNSSVITYPDAQLVPGINGK 69  
 DB 597 SYKNLDGVNDEEDIDVILKSTIILNDIINDIISDISGNSSVITYPDAQLVPGINGK 656  
 QY 70 AILNNSESEVIVKAMDIENDMFNFTVSFWLRPKVSAHLEQYGTNEYSIISSMK 129  
 DB 657 AILNNSESEVIVKAMDIENDMFNFTVSFWLRPKVSAHLEQYGTNEYSIISSMK 716  
 QY 130 KHSLSGWSVSLKGNLILWTLKDSAGEVROITPDLDPKFNAYLANKWVFITINDRL 189  
 DB 717 KHSLSGWSVSLKGNLILWTLKDSAGEVROITPDLDPKFNAYLANKWVFITINDRL 776  
 QY 190 SSANLYINGVLMGSAEITGLAIRREDNNTTLKLDRCNNNNQYVSIDKFRIFCKALNPKIE 249  
 DB 777 SSANLYINGVLMGSAEITGLAIRREDNNTTLKLDRCNNNNQYVSIDKFRIFCKALNPKIE 836  
 QY 250 EKLTSYLSITFLRDFWGNPLKRYDTREYLLIPVASSSKDVOLKNTTDMYLTNAPSTYNGK 309  
 DB 837 EKLTSYLSITFLRDFWGNPLKRYDTREYLLIPVASSSKDVOLKNTTDMYLTNAPSTYNGK 896  
 QY 310 LNIYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGPKDGNAFNNL 369  
 DB 897 LNIYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGPKDGNAFNNL 956  
 QY 370 IRLRFGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNR 429  
 DB 957 IRLRFGYNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNGQIGNDPNR 1016  
 QY 430 DILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 463  
 DB 1017 DILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 1050

RESULT 9  
 AAE07902  
 ID AAE07902 standard; Protein; 1112 AA.  
 AC AAE07902;  
 DT 01-NOV-2001 (first entry)  
 DE C. botulinum C2 translocation domain with TeNT binding domain #1.  
 KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;  
 KM tumour; infection; neurodegenerative disease; gene therapy;  
 KW botulinum neurotoxin; tetanus neurotoxin; TeNT.  
 XX OS Clostridium botulinum.

OS Clostridium tetani.  
 XX WO200158936-A2.  
 XX 16-AUG-2001.  
 XX 04-DEC-2000; 2000MO-GB04644.  
 XX 02-DEC-1999; 99GB-0028530.  
 PR 07-APR-2000; 2000GB-0008658.  
 XX  
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
 PI Shone CC, Sutton JM, Silman N;  
 DR WPI: 2001-514643/56.  
 XX  
 PT New non toxic polypeptide for delivery of a therapeutic agent for the  
 PT treatment of a CNS disorder comprising a binding domain that  
 PT translocates the therapeutic agent into the neuronal cells -  
 XX  
 PS Example 2; Page 49; 50pp; English.

The invention relates to a non toxic polypeptide, for delivery of a  
 CC therapeutic agent to a neuronal cell, which comprises a binding domain  
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated  
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino  
 CC terminal half of HC, designated as HN), that translocates the therapeutic  
 CC agent into the neuronal cell, where the translocation domain is not a HN  
 CC a HN domain of a clostridial neurotoxin and is not a fragment or derivative of  
 CC a useful for the treatment of a disease state associated with neuronal  
 CC cells. The polypeptide constructs are useful for delivering therapeutic  
 CC substances to neuronal cells. They are useful for treating disorders of the  
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours  
 CC and infection. They are also useful in gene therapy. The present sequence  
 CC is C. botulinum C2 enterotoxin translocation domain with tetanus  
 CC neurotoxin (Tent) binding domain used in the exemplification of the  
 CC invention.

Sequence 1112 AA;

Query Match 97.6%; Score 2397; DB 22; Length 1112;  
 Best Local Similarity 99.6%; Pred. No. 2.2e-185;  
 Matches 451; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 YSKNDQWVNEEDIDVILKSTIILNDINNDISIGFSSVITYPDAQVPGINGKA 70  
 DB 658 FTKNDQWVNEEDIDVILKSTIILNDINNDISIGFSSVITYPDAQVPGINGKA 717  
 QY 71 IHLVNNSESEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQGTNEYSIISSMKK 130  
 DB 718 IHLVNNSESEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQGTNEYSIISSMKK 777  
 QY 131 HSLISIGSGVSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKWVEITITNDRLS 190  
 DB 778 HSLISIGSGVSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKWVEITITNDRLS 837  
 QY 191 SANLYINGVLMGSAEITGGAIRREDNNITLKLDRCNNNOYVSIIDKFRIFCKALNPKREIE 250  
 DB 838 SANLYINGVLMGSAEITGGAIRREDNNITLKLDRCNNNOYVSIIDKFRIFCKALNPKREIE 897  
 QY 251 KLYTSTYLTFTFRDQWGNPLRYDTEYLLIPVASSSKDVOLEKNTIDMYLTNAPSYNGKL 310  
 DB 898 KLYTSTYLTFTFRDQWGNPLRYDTEYLLIPVASSSKDVOLEKNTIDMYLTNAPSYNGKL 957  
 QY 311 NIYYRRLNGLKFIIKRYTPNNEIDSFVKSQDFIKLYSYNNNEHTVGPCKGNANFNND 370  
 DB 958 NIYYRRLNGLKFIIKRYTPNNEIDSFVKSQDFIKLYSYNNNEHTVGPCKGNANFNND 1017  
 QY 371 RILRVGYNAPGIIPLVKKMEAVKLRDLKTYVQVLKLYDDKNASLGIVGTHNGOIGPNRND 430  
 DB 1018 RILRVGYNAPGIIPLVKKMEAVKLRDLKTYVQVLKLYDDKNASLGIVGTHNGOIGPNRND 1077

QY 431 ILIASWYFNHLKDKITLGGDWYFVPTDEGWTND 463  
 DB 1078 ILIASWYFNHLKDKITLGGDWYFVPTDEGWTND 1110

RESULT 10  
 AAR12471  
 ID AAR12471 standard; Protein; 452 AA.  
 XX  
 AC AAR12471;  
 XX  
 DT 05-AUG-1991 (first entry)  
 XX  
 DE Tetanus toxin fragment C encoded by gene with increased G+C content.  
 XX  
 KM Terminator; vaccine.  
 XX  
 OS Synthetic.  
 OS  
 PN EP430645-A.  
 XX  
 PD 05-JUN-1991.  
 XX  
 PE 27-NOV-1990; 90EP-0312870.  
 PE  
 PR 17-MAR-1990; 90GB-0006097.  
 PR 28-NOV-1989; 89GB-0026832.  
 XX  
 PA (WELL ) WELLCOME FOUNDATION LTD.  
 PI Makoff AJ, Romanos MA, Clare JJ, Fairweather NF;  
 PI WPI: 1991-166115/23.  
 DR N-PSDB; AAO12121.  
 DR

DNA sequence encoding tetanus toxin fragment C - useful in the  
 PT manufacture of vaccines for immunity to tetanus utilising yeast  
 PT as host organism.  
 XX  
 PS Disclosure; Fig 2; 50pp; English.

The (G+C) content of the synthetic gene is increased by 47% wrt the  
 CC native sequence. This eliminates six "terminator" regions which  
 CC were found to be present in (A+T) rich regions. The terminators  
 CC (termination/endo-nucleolytic processing/polyadenylation sites)  
 CC were previously responsible for incomplete transcription of the  
 CC mRNA. The elimination of these elements (using codon degeneracy)  
 CC provided for successful expression in yeast of the tetanus toxin  
 CC fragment C.

Sequence 452 AA;

Query Match 97.4%; Score 2393; DB 12; Length 452;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-185;  
 Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDQWVNEEDIDVILKSTIILNDINNDISIGFSSVITYPDAQVPGINGKA 72  
 DB 2 KNDQWVNEEDIDVILKSTIILNDINNDISIGFSSVITYPDAQVPGINGKA 61  
 QY 73 LVNNSESEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQGTNEYSIISSMKKHS 132  
 DB 62 LVNNSESEVIVHKAMDIEYNDMFNFVFWLRVPKVSASHLEQGTNEYSIISSMKKHS 121  
 QY 133 LSTIGSGVSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKWVEITITNDRLSSA 192  
 DB 122 LSTIGSGVSVSLKGNLITWLKDSAGEVROITFRDLPKFNAYLANKWVEITITNDRLSSA 181  
 QY 193 NLVYNGVLMGSAEITGGAIRREDNNITLKLDRCNNNOYVSIIDKFRIFCKALNPKREIE 252  
 DB 182 NLVYNGVLMGSAEITGGAIRREDNNITLKLDRCNNNOYVSIIDKFRIFCKALNPKREIE 241

```

OY 253 YTSYLSTFLRDFMGNPLRYDTEYYLIIPVASSSKDYOLKNTIDYMYLTNAPSYTNGKLN1 312
    |||||||
DB 242 YTSYLSTFLRDFMGNPLRYDTEYYLIIPVASSSKDYOLKNTIDYMYLTNAPSYTNGKLN1 301
    |||||||
OY 313 YVRLYLNGIKFLIKRYTPNNEIDSFVKSDFIKLYVSYNNNEHIVGYPKDGNAFNMLDRI 372
    |||||||
DB 302 YVRLYLNGIKFLIKRYTPNNEIDSFVKSDFIKLYVSYNNNEHIVGYPKDGNAFNMLDRI 361
    |||||||
OY 373 LRVGNAGGIPLYKMEAVKLRDLKTSYVOLKLYDPKNAISLGLVTHNGOIGNDPNDIL 432
    |||||||
DB 362 LRVGNAGGIPLYKMEAVKLRDLKTSYVOLKLYDPKNAISLGLVTHNGOIGNDPNDIL 421
    |||||||
OY 433 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 463
    |||||||
DB 422 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 452
    |||||||

RESULT 11
AAB31429 standard; protein: 661 AA.
XX
AC AAB31429;
XX
DT 26-APR-2001 (first entry)
XX
DE Shed antigen-specific B cell antigen linked to tetanus toxin fragment C.
XX
KW Vaccine; shed antigen-specific B cell; idiotypic antibody.
XX
KW Immune complex-mediated disease; autoimmune disease; tetanus protein;
    humoral immune response; cancer.
XX
OS Synthetic.
XX
OS Clostridium tetani.
XX
PN W0200076319-A1.
XX
PD 21-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16677.
XX
PR 16-JUN-1999; 1999US-0139521.
XX
PR 15-JUN-2000; 2000US-0594985.
XX
PA (BIOC-) BIOCRYSTAL LTD.
XX
PI Nelson MB, Barbera-Guillem E;
XX
WPI: 2001-080635/09.
XX
DR
XX
PT Inducing an immune response against shed antigen-specific B cell
    polynucleotides encoding an idiotypic determinant or peptides
    comprising an idiotypic determinant.
XX
PS Example 2; Page 73-76; 81pp; English.
XX
CC The present sequence represents a fusion protein, comprising a protein
    used for immunising against shed antigen-specific B cells linked to a
    fragment of tetanus protein. It is used in vaccines of the invention. The
    specification describes a method for inducing an immune response reactive
    with idiotypes present on shed antigen-specific B cells (SAB) of an
    individual. The method involves administering a vaccine formulation
    comprising polynucleotide encoding an idiotype of an antibody that binds
    to an epitope of shed antigen. The method is useful for inducing an
    immune response reactive with idiotypes present on SAB of an individual.
    The method is useful for depleting shed antigen-specific B cells, and
    for treating immune complex-mediated disease progression in organ
    specific autoimmune disease exacerbated by humoral immune response
    against groups expressed on shed antigen, or by plasma cell production
    of antibodies against groups of shed antigen. It is useful in cancer
    therapy and for treating autoimmune disease.
XX
SQ Sequence 661 AA;

```

```

Query Match 97.2%; Score 2387.5; DB 22; Length 661;
Best Local Similarity 95.2%; Pred. No. 6e-185;
Matches 453; Conservative 2; Mismatches 2; Indels 19; Gaps 1;

OY 7 IPEYS-----KNIDCWNEEDIDVILKSTILLNDINDIISDI 47
    :|:|
DB 186 VPFRRSGSGVWDLFFSHNPNGGLEKNIDCWVNEEDIDVILKSTIILLNDINDIISDI 245
    |||||||
OY 48 SGFNSSVITTPDAQVGLVGINCKALHIVNNESEVIVHKAAMDIEYMDMNFVSPFLRVP 107
    |||||||
DB 246 SGFNSSVITTPDAQVGLVGINCKALHIVNNESEVIVHKAAMDIEYMDMNFVSPFLRVP 305
    |||||||
OY 108 KVSASHLEOYGTNNEYSIISMKKHSLSIGSGSVSLKGNLIMTLKDSAGEVROITFRDL 167
    |||||||
DB 306 KVSASHLEOYGTNNEYSIISMKKHSLSIGSGSVSLKGNLIMTLKDSAGEVROITFRDL 365
    |||||||
OY 168 PDKFNAYLANKWFVITTTNDRILSSANLYINGVLMGSAEITIGAIREDNNITLKDRCNN 227
    |||||||
DB 366 PDKFNAYLANKWFVITTTNDRILSSANLYINGVLMGSAEITIGAIREDNNITLKDRCNN 425
    |||||||
OY 228 NNOYVSIDKFRIFCKALNPKETIEKLYSTFLRDFMGNPLRYDTEYYLIIPVASSSKD 287
    |||||||
DB 426 NNOYVSIDKFRIFCKALNPKETIEKLYSTFLRDFMGNPLRYDTEYYLIIPVASSSKD 485
    |||||||
OY 288 VOLKNTIDYMYLTNAPSYTNGKLNLYRRLYNGIKFLIKRYTPNNEIDSFVKSDFIKLY 347
    |||||||
DB 486 VOLKNTIDYMYLTNAPSYTNGKLNLYRRLYNGIKFLIKRYTPNNEIDSFVKSDFIKLY 545
    |||||||
OY 348 VSYNNNEHIVGYPKDGNAFNMLDRIILRVGYNAPGIPLYKKMAVYLRDLKTSYVOLKLYD 407
    |||||||
DB 546 VSYNNNEHIVGYPKDGNAFNMLDRIILRVGYNAPGIPLYKKMAVYLRDLKTSYVOLKLYD 605
    |||||||
OY 408 DKNASLGLVTHNGOIGNDPNDILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 463
    |||||||
DB 606 DKNASLGLVTHNGOIGNDPNDILIASNMYFNHLKDKILGCDWYFVPTDEGWTND 661
    |||||||

RESULT 12
AAB31427
ID AAB31427 standard; protein: 453 AA.
XX
AC AAB31427;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of tetanus toxin fragment C.
XX
KW Vaccine; shed antigen-specific B cell; idiotypic antibody.
XX
KW Immune complex-mediated disease; autoimmune disease; tetanus protein;
    humoral immune response; cancer.
XX
OS Clostridium tetani.
XX
PN W0200076319-A1.
XX
PD 21-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16677.
XX
PR 16-JUN-1999; 99US-0139521.
XX
PR 15-JUN-2000; 2000US-0594985.
XX
PA (BIOC-) BIOCRYSTAL LTD.
XX
PI Nelson MB, Barbera-Guillem E;
XX
WPI: 2001-080635/09.
XX
DR
XX
PT Inducing an immune response against shed antigen-specific B cell
    polynucleotides encoding an idiotypic determinant or peptides
    comprising an idiotypic determinant.
XX
SQ Sequence 453 AA;

```



XX Example 2; Page 72-73; 81pp; English.

PS The present sequence represents a fragment of tetanus protein, which

CC is used as an immunostimulatory protein in vaccines of the invention. The

CC specification describes a method for inducing an immune response reactive

CC with idiotypes present on shed antigen-specific B cells (SAB) of an

CC individual. The method involves administering a vaccine formulation

CC comprising polynucleotide encoding an idiotypic of an antibody that binds

CC to an epitope of shed antigen. The method is useful for inducing an

CC immune response reactive with idiotypes present on SAB of an individual.

CC The method is useful for depleting shed antigen-specific B cells, and

CC for treating immune complex-mediated disease progression in organ

CC specific autoimmune disease exacerbated by humoral immune response

CC against groups expressed on shed antigen, or by plasma cell production

CC of antibodies against groups of shed antigen. It is useful in cancer

CC therapy and for treating autoimmune disease.

XX Sequence 453 AA:

Query Match 97.1%; Score 2385; DB 22; Length 453;

Best Local Similarity 99.8%; Pred. No. 5.5e-185;

Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 KNIDCWVNDIEDIDVILKKSITLNDINNDIISDGSNSVITPPAQLYVGINGKAH 72

DB 3 KNIDCWVNDIEDIDVILKKSITLNDINNDIISDGSNSVITPPAQLYVGINGKAH 62

QY 73 LVNSESSEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQYGTNEYSLISSMKKHS 132

DB 63 LVNSESSEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQYGTNEYSLISSMKKHS 122

QY 133 LSTGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVITITNDRLSSA 192

DB 123 LSTGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVITITNDRLSSA 182

QY 193 NLYINGVLMGSAEITIGALREDNNITLKDRCNNNNNOYVSIDKFRIFCKALNPEIEKLT 252

DB 183 NLYINGVLMGSAEITIGALREDNNITLKDRCNNNNNOYVSIDKFRIFCKALNPEIEKLT 242

QY 253 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLEKNTDWMYLTNAPSYNGKLN 312

DB 243 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLEKNTDWMYLTNAPSYNGKLN 302

QY 313 YTRRLYNGLKFIIRKRYTPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 372

DB 303 YTRRLYNGLKFIIRKRYTPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 362

QY 373 LRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDIL 432

DB 363 LRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDIL 422

QY 433 IASNMVFNHLKDKILGCDWYFVPTDEGWTND 463

DB 423 IASNMVFNHLKDKILGCDWYFVPTDEGWTND 453

RESULT 13

AAE07897 standard; Protein; 605 AA.

AC AAE07897;

XX 01-NOV-2001 (first entry)

DE Modified clostridial heavy chain fragment #4.

XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;

KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;

KM diphtheria neurotoxin; tetanus neurotoxin; TeNT.

XX Chimeric - Corynebacterium diphtheriae.

OS Chimeric - Clostridium tetani.

XX WO200158936-A2.

PN 16-AUG-2001.

XX 04-DEC-2000; 2000WO-GB04644.

XX 02-DEC-1999; 99GB-0028530.

XX 07-APR-2000; 2000GB-0008658.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Shone CC, Sutton JM, Silman N;

XX WPI, 2001-514643/56.

DR New non toxic polypeptide for delivery of a therapeutic agent for the

PT treatment of a CNS disorder comprising a binding domain that

XX translocates the therapeutic agent into the neuronal cells -

PS Example 2; Page 45; 50pp; English.

CC The invention relates to a non toxic polypeptide, for delivery of a

CC therapeutic agent to a neuronal cell, which comprises a binding domain

CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated

CC as HC) that binds to the neuronal cell and a translocation domain (amino

CC terminal half of HC, designated as HN), that translocates the therapeutic

CC agent into the neuronal cell, where the translocation domain is not a HN

CC domain of a clostridial neurotoxin and is not a fragment or derivative of

CC a HN domain of a clostridial toxin. Polypeptides of the invention are

CC useful for the treatment of a disease state associated with neuronal

CC cells. The polypeptide constructs are useful for delivering therapeutic

CC substances to neuronal cells. They are useful to treat disorders of the

CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours

CC and infection. They are also useful in gene therapy. The present sequence

CC is modified clostridial heavy chain fragment. This sequence is

CC constructed by fusing the truncated binding domain of tetanus neurotoxin

CC (TeNT) with translocation domain of diphtheria neurotoxin.

XX Sequence 605 AA:

Query Match 84.4%; Score 2073; DB 22; Length 605;

Best Local Similarity 100.0%; Pred. No. 1.7e-159;

Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 NNESESEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQYGTNEYSLISSMKKHSLS 134

DB 215 NNESESEVIVHKAMDIEYNDMENNFTVSEFWLRVPKVSASHLEQYGTNEYSLISSMKKHSLS 124

QY 135 IGSWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVITITNDRLSSANL 194

DB 275 IGSWSVSLKGNLIMTLKDSAGEVROITFRDLPKFNAYLANKWVITITNDRLSSANL 184

QY 195 YINGVLMGSAEITIGALREDNNITLKDRCNNNNNOYVSIDKFRIFCKALNPEIEKLT 254

DB 335 YINGVLMGSAEITIGALREDNNITLKDRCNNNNNOYVSIDKFRIFCKALNPEIEKLT 244

QY 255 SYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLEKNTDWMYLTNAPSYNGKLN 314

DB 395 SYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLEKNTDWMYLTNAPSYNGKLN 304

QY 315 RRLYNGLKFIIRKRYTPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 374

DB 455 RRLYNGLKFIIRKRYTPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPKDGNAPNNDRI 364

QY 375 VGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDIL 434

DB 515 VGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGIVGTHNGOIGNDPNRDIL 424

QY 435 SNMVFNHLKDKILGCDWYFVPTDEGWTND 463

DB 575 SNMVFNHLKDKILGCDWYFVPTDEGWTND 453



CC substances to neuronal cells. They are useful to treat disorders of the  
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours  
CC and infection. They are also useful in gene therapy. The present sequence  
CC is modified clostridial heavy chain fragment. This sequence is  
CC constructed by fusing the binding domain II of tetanus neurotoxin (TnT)  
CC with translocation domain of diphtheria neurotoxin.  
XX

SQ Sequence 448 AA;

Query Match 51.3%; Score 1261; DB 22; Length 448;  
Best Local Similarity 100.0%; Pred. No. 8e-94;

Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 YVSIDKFRIFCKALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLIPIVASSSKDVQL 290  
|||  
Db 214 YVSIDKFRIFCKALNPKIEKLYTSYSTFLRDFWGNPLRYDTEYLIPIVASSSKDVQL 273  
QY 291 KNTDYMVLNAPSYTNGKINYYRRLYNGLKFTIKRTPNNEIDSFVKSQDFIKLYVS 350  
|||  
Db 274 KNTDYMVLNAPSYTNGKINYYRRLYNGLKFTIKRTPNNEIDSFVKSQDFIKLYVS 333  
QY 351 NNEHIVGYPKDGNAFNNLDRIIRVGNAPGIPLYKKMEAVKLRDLKTYVOLKLYDDKN 410  
|||  
Db 334 NNEHIVGYPKDGNAFNNLDRIIRVGNAPGIPLYKKMEAVKLRDLKTYVOLKLYDDKN 393  
QY 411 ASIGLVGTHNGQIGNDPNRDILLASNMYFNHLKDKILGCDWYFVPTDEGWTND 463  
|||  
Db 394 ASIGLVGTHNGQIGNDPNRDILLASNMYFNHLKDKILGCDWYFVPTDEGWTND 446

Search completed: December 1, 2002, 11:21:13  
Job time : 70 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:20:37 ; Search time 26 Seconds

(Without alignments)  
523.954 Million cell updates/sec

Title: US-09-816-467-2

Perfect score: 2456 1 MFESTPIPFYSKNDICWD.....DKILGDMYFVPTDEGTND 463

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2.6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2.6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2451	99.8	US-08-913-880C-17	Sequence 17, Appl
2	2451	99.8	US-08-913-880C-16	Sequence 16, Appl
3	2451	99.8	US-08-913-880C-15	Sequence 15, Appl
4	2451	99.8	US-08-913-880C-14	Sequence 14, Appl
5	2451	99.8	US-08-913-880C-13	Sequence 13, Appl
6	2451	99.8	US-08-913-880C-12	Sequence 12, Appl
7	2451	99.8	US-08-913-880C-11	Sequence 11, Appl
8	2451	99.8	US-08-913-880C-10	Sequence 10, Appl
9	2451	99.8	US-08-913-880C-1	Sequence 5, Appl
10	2440	99.3	US-07-618-312A-5	Sequence 4, Appl
11	2393	97.4	US-07-618-312A-4	Sequence 2, Appl
12	2380	96.9	US-07-618-312A-2	Sequence 2, Appl
13	2380	96.9	US-07-618-312A-2	Sequence 2, Appl
14	2380	96.9	US-08-280-228-2	Sequence 8, Appl
15	2345	95.5	US-08-110-786A-8	Sequence 28, Appl
16	653.5	26.6	US-08-480-604A-28	Sequence 28, Appl
17	653.5	26.6	US-08-480-604A-28	Sequence 28, Appl
18	653.5	26.6	US-08-480-604A-28	Sequence 28, Appl
19	627.5	25.5	US-08-480-604A-23	Sequence 23, Appl
20	627.5	25.5	US-08-480-604A-23	Sequence 23, Appl
21	627.5	25.5	US-08-480-604A-23	Sequence 23, Appl
22	627.5	25.5	US-08-480-604A-26	Sequence 26, Appl
23	627.5	25.5	US-08-480-604A-26	Sequence 26, Appl
24	627.5	25.5	US-08-480-604A-26	Sequence 26, Appl
25	478	19.5	US-09-255-829-20	Sequence 20, Appl
26	220	9.0	US-08-110-786A-5	Sequence 5, Appl
27	168	6.8	PCT-US93-11703-64	Sequence 64, Appl

28	163	6.6	140	4	US-08-446-114A-22	Sequence 22, Appl
29	127	5.2	2391	2	US-08-446-855A-2	Sequence 2, Appl
30	127	5.2	2391	4	US-09-150-741-2	Sequence 2, Appl
31	125	5.1	20	1	US-08-231-437-1	Sequence 1, Appl
32	125	5.1	20	5	PCT-US93-06175-1	Sequence 1, Appl
33	125	5.1	21	1	US-08-231-437-2	Sequence 2, Appl
34	125	5.1	21	1	PCT-US93-06175-2	Sequence 2, Appl
35	119.5	4.9	2366	2	US-08-480-604A-10	Sequence 10, Appl
36	119.5	4.9	2366	2	US-08-480-604A-10	Sequence 10, Appl
37	119.5	4.9	2366	4	US-08-915-136-10	Sequence 10, Appl
38	119.5	4.9	2366	4	US-08-915-136-10	Sequence 10, Appl
39	115	4.7	993	4	US-08-836-687B-30	Sequence 30, Appl
40	114	4.6	32	1	US-08-446-692-14	Sequence 14, Appl
41	114	4.6	32	2	US-08-446-692-14	Sequence 14, Appl
42	112.5	4.6	1005	4	US-09-206-942-41	Sequence 41, Appl
43	112.5	4.6	1011	4	US-09-206-942-39	Sequence 39, Appl
44	112	4.6	21	1	US-07-610-525-1	Sequence 1, Appl
45	112	4.6	21	2	US-08-661-052-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-08-913-880C-17
: Sequence 17, Application US/08913880C
: Patent No. 6372225
: GENERAL INFORMATION:
:   APPLICANT: MATSUDA, Morihito
:   TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
:   FILE REFERENCE: 216-380P
:   CURRENT APPLICATION NUMBER: US/08/913, 880C
:   NUMBER OF SEQ ID NOS: 17
:   SEQ ID NO 17
:   LENGTH: 853
:   TYPE: PRT
:   ORGANISM: Clostridium tetani
:   FEATURE:
:   OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
US-08-913-880C-17

Query Match          99.8%: Score 2451; DB 4; Length 853;
Best Local Similarity 100.0%: Pred. No. 2.7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VFSTPIPFYSKNDICWDNEEDIDVILKSTIINLDINNDIISDGFNSVITYPPDAQ 61
   |||||
DB 392 VFSTPIPFYSKNDICWDNEEDIDVILKSTIINLDINNDIISDGFNSVITYPPDAQ 451
   |||||
QY 62 LVPGINGKAIHIVNSESSEVIVHKAMDEYDMRNFTVSWLRPKVYSASHLEQYGTNE 121
   |||||
DB 452 LVPGINGKAIHIVNSESSEVIVHKAMDEYDMRNFTVSWLRPKVYSASHLEQYGTNE 511
   |||||
QY 122 YSIISMKKHSLSIGSGWSYSLKGNLITLKDAGSEVROITFRDLDPKFNAYLANKKWF 181
   |||||
DB 512 YSIISMKKHSLSIGSGWSYSLKGNLITLKDAGSEVROITFRDLDPKFNAYLANKKWF 571
   |||||
QY 182 ITTNRRLSANIYINGVLMGSAEITGLGAIREDNNITLTKIDRCNNNNQVYSIDKFRIFC 241
   |||||
DB 572 ITTNRRLSANIYINGVLMGSAEITGLGAIREDNNITLTKIDRCNNNNQVYSIDKFRIFC 631
   |||||
QY 242 KALNPKIEIKLYSYLSTIFLRDFWGNPLRYDPEYILIPVASSSKDVQLKITTYMYLTN 301
   |||||
DB 632 KALNPKIEIKLYSYLSTIFLRDFWGNPLRYDPEYILIPVASSSKDVQLKITTYMYLTN 691
   |||||
QY 302 AASYTNGKLTNYYRRLYNGIKFKIKRYTPNNEIDSFKVSGDFIKLYSYNNNEHIVGYPK 361
   |||||
DB 692 AASYTNGKLTNYYRRLYNGIKFKIKRYTPNNEIDSFKVSGDFIKLYSYNNNEHIVGYPK 751
   |||||
QY 362 DGNAFNNLDRITLRYGVNAPGIPLYKKMEAKVLDLKITYSVQLKYIDKRNASLGLVGHNG 421
   |||||
```

```

Db 752 DGNMFNNDRILRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDXKSLGLVGTNG 811
Qy 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 812 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 853

RESULT 2
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 16
; LENGTH: 858
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
US-08-913-880C-16

Query Match          99.8%; Score 2451; DB 4; Length 858;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VFSTPIPSYSKNLDCWVNEEDIDYILKSTILNDINNDISDSIGFNSVITYPPDAQ 61
Db 397 VFSTPIPSYSKNLDCWVNEEDIDYILKSTILNDINNDISDSIGFNSVITYPPDAQ 456
Qy 62 LVPINGKAIHLVNNSESEYIVHKADIEYNDMFNFTVSFWLRVPKVSASHLEOYGTNE 121
Db 457 LVPINGKAIHLVNNSESEYIVHKADIEYNDMFNFTVSFWLRVPKVSASHLEOYGTNE 516
Qy 122 YSISSMKRHSISIGSGWSYSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWF 181
Db 517 YSISSMKRHSISIGSGWSYSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWF 576
Qy 182 ITTNDRLSSANLYINGVLMSAEITGLGAIREDNNITTLKDRCNNNOQVYSIDKFRIFC 241
Db 577 ITTNDRLSSANLYINGVLMSAEITGLGAIREDNNITTLKDRCNNNOQVYSIDKFRIFC 636
Qy 242 KALNPKIEIKLYTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMTLN 301
Db 637 KALNPKIEIKLYTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMTLN 696
Qy 302 APSYTNKLNLYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 697 APSYTNKLNLYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 756
Qy 362 DGNMFNNDRILRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDDKNASLGLVGTNG 421
Db 757 DGNMFNNDRILRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDDKNASLGLVGTNG 816
Qy 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 817 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 858

RESULT 3
US-08-913-880C-15
; Sequence 15, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P

```

```

; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 15
; LENGTH: 860
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1
US-08-913-880C-15

Query Match          99.8%; Score 2451; DB 4; Length 860;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VFSTPIPSYSKNLDCWVNEEDIDYILKSTILNDINNDISDSIGFNSVITYPPDAQ 61
Db 397 VFSTPIPSYSKNLDCWVNEEDIDYILKSTILNDINNDISDSIGFNSVITYPPDAQ 458
Qy 62 LVPINGKAIHLVNNSESEYIVHKADIEYNDMFNFTVSFWLRVPKVSASHLEOYGTNE 121
Db 457 LVPINGKAIHLVNNSESEYIVHKADIEYNDMFNFTVSFWLRVPKVSASHLEOYGTNE 518
Qy 122 YSISSMKRHSISIGSGWSYSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWF 181
Db 519 YSISSMKRHSISIGSGWSYSLKGNLIMTLKDSAGEVROITFRDLPDKFNAYLANKWF 578
Qy 182 ITTNDRLSSANLYINGVLMSAEITGLGAIREDNNITTLKDRCNNNOQVYSIDKFRIFC 241
Db 579 ITTNDRLSSANLYINGVLMSAEITGLGAIREDNNITTLKDRCNNNOQVYSIDKFRIFC 638
Qy 242 KALNPKIEIKLYTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMTLN 301
Db 639 KALNPKIEIKLYTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNITDYMTLN 698
Qy 302 APSYTNKLNLYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 699 APSYTNKLNLYRRLYNGLKFTIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 758
Qy 362 DGNMFNNDRILRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDDKNASLGLVGTNG 421
Db 759 DGNMFNNDRILRVGNAPGIPLYKKMEAVKLRDLKTSVOLKLYDDKNASLGLVGTNG 818
Qy 422 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 819 QIGNDPNRDLIASNMWYFNHLKDKILGCDWYFVPTDEGWTND 860

RESULT 4
US-08-913-880C-14
; Sequence 14, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiko
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
; LENGTH: 862
; TYPE: PRP
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1
US-08-913-880C-14

Query Match          99.8%; Score 2451; DB 4; Length 862;
Best Local Similarity 100.0%; Pred. No. 2,7e-189;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VFSTPIPSYSKNLDCWVNEEDIDYILKSTILNDINNDISDSIGFNSVITYPPDAQ 61

```

```
|||||
Db 401 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDISGNSSVITYPDAQ 460
QY 62 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 121
Db 461 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 520
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
Db 521 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 580
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241
Db 581 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 640
QY 242 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 301
Db 641 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 700
QY 302 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 701 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 760
QY 362 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 761 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 820
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 821 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 862
```

RESULT 5  
US-08-913-880C-13

; Sequence 13, Application US/08913880C  
; Patent No. 6372225

; GENERAL INFORMATION:

; APPLICANT: MATSUDA, MORIHIRO

; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS

; FILE REFERENCE: 216-380P

; CURRENT APPLICATION NUMBER: US/08/913,880C

; CURRENT FILING DATE: 1997-11-24

; NUMBER OF SEQ ID NOS: 17

; SEQ ID NO 13

; LENGTH: 865

; TYPE: PRT

; ORGANISM: Clostridium tetani

; FEATURE:

; OTHER INFORMATION: Amino Acids 451 to 1315 of SEQ ID NO: 1

US-08-913-880C-13

```
Query Match 99.8%; Score 2451; DB 4; Length 865;  
Best Local Similarity 100.0%; Pred. No. 2,7e-189;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDISGNSSVITYPDAQ 61  
Db 404 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDISGNSSVITYPDAQ 463  
QY 62 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 121  
Db 464 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 523  
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 181  
Db 524 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 583  
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241  
Db 584 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 643  
QY 242 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 301
```

```
|||||
Db 644 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 703
QY 302 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 704 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 763
QY 362 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 764 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 823
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 824 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 865
```

RESULT 6  
US-08-913-880C-12

; Sequence 12, Application US/08913880C  
; Patent No. 6372225

; GENERAL INFORMATION:

; APPLICANT: MATSUDA, MORIHIRO

; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS

; FILE REFERENCE: 216-380P

; CURRENT APPLICATION NUMBER: US/08/913,880C

; CURRENT FILING DATE: 1997-11-24

; NUMBER OF SEQ ID NOS: 17

; SEQ ID NO 12

; LENGTH: 866

; TYPE: PRT

; ORGANISM: Clostridium tetani

; FEATURE:

; OTHER INFORMATION: Amino Acids 450 to 1315 of SEQ ID NO: 1

US-08-913-880C-12

Query Match 99.8%; Score 2451; DB 4; Length 866;  
Best Local Similarity 100.0%; Pred. No. 2,7e-189;  
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDISGNSSVITYPDAQ 61
Db 405 VFSPIPFYSKNDLCWVNEEDIDVILKSTILNDINNDIISDISGNSSVITYPDAQ 464
QY 62 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 121
Db 465 LVPINGKAHLVNNESSEVIYHKAMDEIYNDMFNFVSWLRPKVSASHLEOYGTNE 524
QY 122 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 181
Db 525 YSIISMKKHSLSIGSGWSVSLKGNLIMWLKDSAGEVROITFRDLPDKFNAYLANKWVF 584
QY 182 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 241
Db 585 ITTNDRLSSANLYINGVLGMSAETTGAIREDNNITLKDRCNNNNQYVSIDKFRIFC 644
QY 242 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 301
Db 645 KALNPKEIEKLYTSLITFLRDFWGNPLRYDTEYLLIPVASSSKDVOQLKNTIDYMYLTN 704
QY 302 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361
Db 705 APSYTNGLKNIYRRLYNGLKFTIKRTPPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 764
QY 362 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421
Db 765 DGNAFNNLDRLIRVGNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 824
QY 422 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463
Db 825 QIGNDPNRDLILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 866
```

RESULT 7  
 US-08-913-880C-11  
 ; Sequence 11, Application US/08913880C  
 ; Patent No. 6372225  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MATSUDA, Morihito  
 ; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
 ; FILE REFERENCE: VACCINE  
 ; CURRENT APPLICATION NUMBER: US/08/913,880C  
 ; CURRENT FILING DATE: 1997-11-24  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SEQ ID NO 11  
 ; LENGTH: 874  
 ; TYPE: PRP  
 ; ORGANISM: Clostridium tetani  
 ; FEATURE:  
 ; OTHER INFORMATION: Amino Acids 442 to 1315 of SEQ ID NO: 1  
 US-08-913-880C-11

Query Match 99.8%; Score 2451; DB 4; Length 874;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-189;  
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61  
 DB 413 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 472  
 QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 121  
 DB 473 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 532  
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVQITFRDLPDKFNAYLANKWVF 181  
 DB 533 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVQITFRDLPDKFNAYLANKWVF 592  
 QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 241  
 DB 593 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 652  
 QY 242 KALNPKIEKLYTSYLTFLRDFWGNPLRYDTEYLLIPVASSSKDVQLKNTIDYMLTN 301  
 DB 653 KALNPKIEKLYTSYLTFLRDFWGNPLRYDTEYLLIPVASSSKDVQLKNTIDYMLTN 712  
 QY 302 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361  
 DB 713 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 772  
 QY 362 DGNAPNNLDRIILRVGNVAPGIPLYKKMEAVKLRDLKTVSVQKLKYDDKNASLGLVGTNG 421  
 DB 773 DGNAPNNLDRIILRVGNVAPGIPLYKKMEAVKLRDLKTVSVQKLKYDDKNASLGLVGTNG 832  
 QY 422 QIGNDPNRDIILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463  
 DB 833 QIGNDPNRDIILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 874

RESULT 8  
 US-08-913-880C-10  
 ; Sequence 10, Application US/08913880C  
 ; Patent No. 6372225  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MATSUDA, Morihito  
 ; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
 ; FILE REFERENCE: VACCINE  
 ; CURRENT APPLICATION NUMBER: US/08/913,880C  
 ; CURRENT FILING DATE: 1997-11-24  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SEQ ID NO 10  
 ; LENGTH: 875  
 ; TYPE: PRP  
 ; ORGANISM: Clostridium tetani

FEATURE:  
 ; OTHER INFORMATION: Amino Acids 441 to 1315 of SEQ ID NO: 1  
 US-08-913-880C-10

Query Match 99.8%; Score 2451; DB 4; Length 875;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-189;  
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61  
 DB 414 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 473  
 QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 121  
 DB 474 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 533  
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVQITFRDLPDKFNAYLANKWVF 181  
 DB 534 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVQITFRDLPDKFNAYLANKWVF 593  
 QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 241  
 DB 594 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNITLKLDRCNNNNQYVSIDKFRIFC 653  
 QY 242 KALNPKIEKLYTSYLTFLRDFWGNPLRYDTEYLLIPVASSSKDVQLKNTIDYMLTN 301  
 DB 654 KALNPKIEKLYTSYLTFLRDFWGNPLRYDTEYLLIPVASSSKDVQLKNTIDYMLTN 713  
 QY 302 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 361  
 DB 714 APSYNGKLNIIYRRLYNGKFLIKRTYPNNEIDSFVKSQDFIKLYSYNNNEHIVGYPK 773  
 QY 362 DGNAPNNLDRIILRVGNVAPGIPLYKKMEAVKLRDLKTVSVQKLKYDDKNASLGLVGTNG 421  
 DB 774 DGNAPNNLDRIILRVGNVAPGIPLYKKMEAVKLRDLKTVSVQKLKYDDKNASLGLVGTNG 833  
 QY 422 QIGNDPNRDIILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 463  
 DB 834 QIGNDPNRDIILASNMWYFNHLKDKILGCDWYFVPTDEGWTND 875

RESULT 9  
 US-08-913-880C-1  
 ; Sequence 1, Application US/08913880C  
 ; Patent No. 6372225  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MATSUDA, Morihito  
 ; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS  
 ; FILE REFERENCE: VACCINE  
 ; CURRENT APPLICATION NUMBER: US/08/913,880C  
 ; CURRENT FILING DATE: 1997-11-24  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SEQ ID NO 1  
 ; LENGTH: 1315  
 ; TYPE: PRP  
 ; ORGANISM: Clostridium tetani  
 US-08-913-880C-1

Query Match 99.8%; Score 2451; DB 4; Length 1315;  
 Best Local Similarity 100.0%; Pred. No. 5e-189;  
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 61  
 DB 854 VSTPIPFYSKRLDCWVNDNEEDIDVILKSTLINDINDIISDISGFSNVITTPDAQ 913  
 QY 62 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 121  
 DB 914 LVPINGKAIHLVNNSESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEOYGTNE 973  
 QY 122 YSIISMKKHSLSIGSGWSVSLKGNLITWTLKDSAGEVQITFRDLPDKFNAYLANKWVF 181



Db 974 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITFRDLDPDKFNAYLANWKWF 1033

QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKIDRCNNNNQVYSIDKFRIFC 241

Db 1034 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKIDRCNNNNQVYSIDKFRIFC 1093

QY 242 KALNPEIEKLYTSYSTIFLRDPWGNPLRYDTEYLIIPVASSSKDVOLKNTIDMYLTN 301

Db 1094 KALNPEIEKLYTSYSTIFLRDPWGNPLRYDTEYLIIPVASSSKDVOLKNTIDMYLTN 1153

QY 302 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 361

Db 1154 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 1213

QY 362 DGNAPNNDRILRYGVNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421

Db 1214 DGNAPNNDRILRYGVNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 1273

QY 422 QIGNDPNRDLIASNWFNHLKDKILGCDWYFVPTDEGWTND 463

Db 1274 QIGNDPNRDLIASNWFNHLKDKILGCDWYFVPTDEGWTND 1315

## RESULT 10

US-08-668-381A-5

; Sequence 5, Application US/08668381A

; Patent No. 5780024

; GENERAL INFORMATION:

; APPLICANT: Brown, Robert H.

; APPLICANT: Fishman, Paul S.

; APPLICANT: Francis, Jonathan W.

; APPLICANT: Hosler, Betsy A.

; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/668,381A

; FILING DATE: 21-JUN-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/000,473

; FILING DATE: 23-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,164

; REFERENCE/DOCKET NUMBER: 00786/269001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 618 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-668-381A-5

Query Match 99.3%; Score 2440; DB 1; Length 618;

Best Local Similarity 99.6%; Pred. No. 1.3e-188;

Matches 460; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSTPLPESYSKNDLDCWVDNEEDIDVILKKSTILNDINDNDISDISGFNSVITTPDAQ 61

Db 157 VSTPLPESYSKNDLDCWVDNEEDIDVILKKSTILNDINDNDISDISGFNSVITTPDAQ 216

QY 62 LYPGNGKAIHLYNNSSSPVIVKAMDIEYNDFNFTYSFWLRVPKVSASHLEQYGTNE 121

Db 217 LYPGNGKAIHLYNNSSSPVIVKAMDIEYNDFNFTYSFWLRVPKVSASHLEQYGTNE 276

QY 122 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITFRDLDPDKFNAYLANWKWF 181

Db 277 YSIISMKKHSLSIGGWSVSLKGNLIMTLKDSAGEVROITFRDLDPDKFNAYLANWKWF 336

QY 182 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKIDRCNNNNQVYSIDKFRIFC 241

Db 337 ITTNDRLSSANLYINGVLMGSAEITGLGAIREDNNTTLKIDRCNNNNQVYSIDKFRIFC 396

QY 242 KALNPEIEKLYTSYSTIFLRDPWGNPLRYDTEYLIIPVASSSKDVOLKNTIDMYLTN 301

Db 397 KALNPEIEKLYTSYSTIFLRDPWGNPLRYDTEYLIIPVASSSKDVOLKNTIDMYLTN 456

QY 302 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 361

Db 457 APSYNGKLNYYRRLYNGLKFTIKRTPNNEIDSFVKSDFIKLYSYNNNEHIVGPK 516

QY 362 DGNAPNNDRILRYGVNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 421

Db 517 DGNAPNNDRILRYGVNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKNASLGLVGTNG 576

QY 422 QIGNDPNRDLIASNWFNHLKDKILGCDWYFVPTDEGWTND 463

Db 577 QIGNDPNRDLIASNWFNHLKDKILGCDWYFVPTDEGWTND 618

## RESULT 11

US-07-618-312A-4

; Sequence 4, Application US/07618312A

; Patent No. 5389540

; GENERAL INFORMATION:

; APPLICANT: Makoff Dr. Andrew J

; APPLICANT: Romanos Dr. Michael A

; APPLICANT: Clare Dr. Jeffrey J

; APPLICANT: Fairweather Dr. Neil F

; TITLE OF INVENTION: VACCINES

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: 14th Floor

; STREET: 2200 Clarendon Boulevard,

; CITY: Arlington,

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22201

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/618,312A

; FILING DATE: 19910516

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 8926832.0

; FILING DATE: 28-NOV-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9006097.1

; FILING DATE: 17-MAR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Crawford Mr. Arthur R

; REGISTRATION NUMBER: 25,327

; REFERENCE/DOCKET NUMBER: 510-51

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 0101 703 8750400

TELEFAX: 0101 703 5253468  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-618-312A-4

Query Match 97.4%; Score 2393; DB 1; Length 452;  
Best Local Similarity 100.0%; Pred. No. 5.1e-185;  
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDGWDNEEDIDVILKKSTLINDINDIISDGFNSSVITPPDAOLVPGINGKAH 72  
DB 2 KNDGWDNEEDIDVILKKSTLINDINDIISDGFNSSVITPPDAOLVPGINGKAH 61  
QY 73 LVNNESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTIISMKKHS 132  
DB 62 LVNNESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTIISMKKHS 121  
QY 133 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANKWVFTITNDRLSA 192  
DB 122 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANKWVFTITNDRLSA 181  
QY 193 NLYINGVLMGSAEITGLAIREDDNITLKDRCNNNOYVSIDKFRIFCKALNPKKEIEKL 252  
DB 182 NLYINGVLMGSAEITGLAIREDDNITLKDRCNNNOYVSIDKFRIFCKALNPKKEIEKL 241  
QY 253 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 312  
DB 242 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 301  
QY 313 YRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNIDRI 372  
DB 302 YRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNIDRI 361  
QY 373 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNAISLGLVGHNGOIGNDPNRDL 432  
DB 362 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNAISLGLVGHNGOIGNDPNRDL 421  
QY 433 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 463  
DB 422 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 452

## RESULT 12

US-08-280-228-4  
Sequence 4, Application US/08280228

Patent No. 5571694  
GENERAL INFORMATION:  
APPLICANT: Makoff Dr, Andrew J  
APPLICANT: Romanos Dr, Michael A  
APPLICANT: Clare Dr, Jeffrey J  
TITLE OF INVENTION: VACCINES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: NIXON & VANDERHYTE P.C.  
STREET: 1100 No. 5571694th Glebe Road  
CITY: Arlington,  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,228  
FILING DATE: 25-JUL-1994

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/618,312  
FILING DATE: 27-NOV-1990  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8926832.0  
FILING DATE: 28-NOV-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9006097.1  
FILING DATE: 17-MAR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilson, Mary J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 117-163  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-280-228-4

Query Match 97.4%; Score 2393; DB 1; Length 452;  
Best Local Similarity 100.0%; Pred. No. 5.1e-185;  
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KNDGWDNEEDIDVILKKSTLINDINDIISDGFNSSVITPPDAOLVPGINGKAH 72  
DB 2 KNDGWDNEEDIDVILKKSTLINDINDIISDGFNSSVITPPDAOLVPGINGKAH 61  
QY 73 LVNNESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTIISMKKHS 132  
DB 62 LVNNESEVIVHKAMDIENDMFNFTVSFWLRVPKVSASHLEQYGTNEYSTIISMKKHS 121  
QY 133 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANKWVFTITNDRLSA 192  
DB 122 LSTGSGWSVSLKGNLITWLKDSAGEVROITPRDLPDKFNAYLANKWVFTITNDRLSA 181  
QY 193 NLYINGVLMGSAEITGLAIREDDNITLKDRCNNNOYVSIDKFRIFCKALNPKKEIEKL 252  
DB 182 NLYINGVLMGSAEITGLAIREDDNITLKDRCNNNOYVSIDKFRIFCKALNPKKEIEKL 241  
QY 253 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 312  
DB 242 YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKDVOLKNTDYMYLTNAPSYTNGKLN 301  
QY 313 YRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNIDRI 372  
DB 302 YRRLYNGLKFIIRKRYTPNNEIDSFYKSGDFIKLYSYNNNEHIVGPKDGNAFNNIDRI 361  
QY 373 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNAISLGLVGHNGOIGNDPNRDL 432  
DB 362 LRVGYNAPGIPLYKKMAVLRDLKTYSVOLKLYDDKNAISLGLVGHNGOIGNDPNRDL 421  
QY 433 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 463  
DB 422 IASNMYFNHLKDKILGCDWYFVPTDEGWTND 452

## RESULT 13

US-07-618-312A-2  
Sequence 2, Application US/07618312A

Patent No. 538540  
GENERAL INFORMATION:  
APPLICANT: Makoff Dr, Andrew J  
APPLICANT: Romanos Dr, Michael A  
APPLICANT: Clare Dr, Jeffrey J  
APPLICANT: Fairweather Dr, Neil F

```

1 TITLE OF INVENTION: VACCINES
2 NUMBER OF SEQUENCES: 13
3 CORRESPONDENCE ADDRESS:
4 ADDRESS: 14th Floor
5 STREET: 2200 Clarendon Boulevard,
6 CITY: Arlington,
7 STATE: Virginia
8 COUNTRY: U.S.A.
9 ZIP: 22201
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-POS
15 SOFTWARE: Patent Release #1.0, Version #1.25
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/07/618,312A
18 FILING DATE: 19910516
19 CLASSIFICATION: 424
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: GB 8926832.0
22 FILING DATE: 28-NOV-1989
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: GB 9006097.1
25 FILING DATE: 17-MAR-1990
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Crawford Mr, Arthur R
28 REGISTRATION NUMBER: 25,327
29 REFERENCE/DOCKET NUMBER: 510-51
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 0101 703 8750400
32 TELEFAX: 0101 703 5253468
33
34 TELEX: 200797 NIKX UR
35 INFORMATION FOR SEQ ID NO: 2:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 452 amino acids
38 TYPE: AMINO ACID
39 TOPOLOGY: linear
40
41 MOLECULE TYPE: protein
42
43 IS-07-618-312A-2

```

QY	193	NLYINGVLMGSAEFTTGGAIREDNNITLKDRCNNNNQYVIDKFRJECALNPKETEKI	252
Db	182	NLYINGVLMGSAEFTTGGAIREDNNITLKDRCNNNNQYVIDKFRJECALNPKETEKI	241
QY	253	YTSYLSITFLURDFWGNPLRDYTEYLLIPVASSSKDQVLKNIITDYMVLTPNAPSTYNGKLI	312
Db	242	YTSYLSITFLURDFWGNPLRDYTEYLLIPVASSSKDQVLKNIITDYMVLTPNAPSTYNGKLI	301
QY	313	YRRRLYNGKFKTIKRRYPNNNIDSEFSKGFIRKLYVSYNNNEHIVGYPKQGNMFNNLDRI	372
Db	302	YRRRLYNGKFKTIKRRYPNNNIDSEFSKGFIRKLYVSYNNNEHIVGYPKQGNMFNNLDRI	361
QY	373	LRVGNAPGIPLYKKMEAVKLRDCKTYSVOLIKLYDDKNASLGIVGTHNGOIGDNPNDIL	422
Db	362	LRVGNAPGIPLYKKMEAVKLRDCKTYSVOLIKLYDDKNASLGIVGTHNGOIGDNPNDIL	421
QY	433	IASNMYFNHLKDKITLGCOWYFVLPDEGNTN	463
Db	422	IASNMYFNHLKDKITLGCOWYFVLPDEGNTN	452

```

RESULT 15
US-08-110-786A-8
; Sequence 8, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; City: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-110-786A-8

```

Matches	444,	Conservative	0;	Mismatches	7;	Indels	0;	Gaps	0;
QY	13	KNLDCWVNEEDIDVILKKTSTLMLDINNOIISIGFSSVTTTPPAOLVPGINGKAH	72						
Db	2	KNLDCWVNEEDIDVILKKSMTLMLDINNOISIGSSNSRITTPPAOLVPGINGKAH	61						
QY	73	LVNNESEVYIYRKADIEYNDMFNNFTVSPFLRYPKYASASHLEQYGTNEYSTISSMKKHS	132						
Db	62	LVNNESEVYIYRKADIEYNDFMFNNFTVSPFLRYPKYASASHLEQYGTNEYSSINSMKKHS	121						
QY	133	LSIGSWSVSLKGNMLIMTLKDSAGEVQOITFRDLPDKFNAYLANKKVVITTTNDRLSA	192						
Db	122	LSIGSWSVSLKGNMLIMTLKDSAGEVQOITFRDLPDKFNAYLANKKVVITTTNDRLSA	161						
QY	193	NLYINGVLWGSAEITLGAIREDNNITLKLDRCNNNNQYVSIKDEFRIEFCALNPKIEKL	252						
Db	182	NLYINGVLWGSAEITLGAKNREDNNITLKLDRCNNNNQYVSNDFRIEFCALNPKIEKL	241						
QY	253	YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKVQULKNTDYMYLTNAPSYNGKLN	312						
Db	242	YTSYLSITFLRDFWGNPLRYDTEYLLIPVASSSKVQULKNTDYMYLTNAPSYNGKLN	301						
QY	313	YYSRIYNGLKFTIKKRTYPNNEIDSFVKSQDFIKLTVSYVNNNEHLYGVYPKDGAENNDRI	372						
Db	302	YYSRIYNGLKFTIKKRTYPNNEIDSFVKSQDFIKLTVSYVNNNEHLYGVYPKDGAENNDRI	361						
QY	373	LRYGYNAPGIPLKMKMEAVKLRDLKTYVQULKLYDDKNAISGLVSTHNGOIGNDPNRDL	432						
Db	362	LRYGYNAPGIPLKMKMEAVKLRDLKTYVQULKLYDDKNAISGLVSTHNGOIGNDPNRDL	421						
QY	433	IASNMYFNHLKDKLIGCDWYFYVPTDGGWTND	463						
Db	422	IASNMYFNHLKDKLIGCDWYFYVPTDGGWTND	452						

Search completed: December 1, 2002, 11:25:31  
Job time : 29 secs

[illegible]

```

QY 229 NOVVIDKRIEFCALNKEIEKLYTSLTFLKDEKGNPLRDEYLLIPVASSKDY 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1063 HRYIWKIFENLEKELNKEIKLDYDQNSGILKDFWGDYQYKPYMLNLDYPNKYV 1122
QY 289 QLK--ITDYATLNAPSYTNGKLNYYR-RLYNGLKFLIRYPNNIDSPKSGDFIK 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1123 DVNNNGIGYWL-KGPRGSVWTTNLYNSSLYRGTKFLIKKYGASGNK-DNIVRNDRYV 1180
QY 346 LYSVANNHEIYGVKDNNAENNDRLRVGYNAPGIPLYKMEAVKLRDLK----- 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1181 INVYKKEY-----RLATNASQAGEKILSLALPVDGNLSQVYVM 1222
QY 398 -----TYSVOLKLYDDKNASIGLVGTHNGOICNDPNRDLILASNNYFNHLK--DKIL 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1223 KSKNDGJTNNCKMNLQNNNGNDIGIFGH-QFNNIK--LVASNNYNRQIERSSRTL 1277
QY 448 GCDMYFVPEDEGM 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1278 GCSWEIIPVDDGM 1290

```

```

RESULT 2
US-09-288-326-9
/ Sequence 9, Application US/09288326
/ Patent No. US20010018049A1
/ GENERAL INFORMATION:
/ APPLICANT: Kel Roger Aoki
/ APPLICANT: George Sachs
/ TITLE OF INVENTION: Method and Compositions for the
/ TITLE OF INVENTION: Treatment of Pancreatitis
/ FILE REFERENCE: 17282
/ CURRENT APPLICATION NUMBER: US/09/288,326
/ CURRENT FILING DATE: 1999-04-08
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 9
/ LENGTH: 425
/ TYPE: PRT
/ ORGANISM: Clostridium botulinum
US-09-288-326-9

```

```

Query Match 25.5%; Score 627.5; DB 10; Length 425;
Best Local Similarity 33.3%; Pred. No. 4.4e-40;
Matches 153; Conservative 84; Mismatches 149; Indels 73; Gaps 16;

```

```

QY 32 STLINDINNDIISDISGNSVITYPDAOLVPGINGKAIHLVNNESSEVIYKAMDIEY 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 TSLINLRYESNHLIDLSRYASKINIGSKVNFDP-IDKNQIQLFNLESSKEVLKNAIVY 63
QY 92 NDMFNNTVSEWLRVPK---VSASHLEQYGTNEYSTISSMKKHSLSISGWSVSLKGN 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 NSMTEENSTSFWRITRPYFNSISLN-----NFTIINCENN-----SGMKVSLNAGE 111
QY 148 LIWTLKDSAGEVROIFTR-----DLDPKFNAVLANKWVFTITNDRLSSANLYINGVLMG 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 112 IIMTLQOTQIEKQVAVKYSOMINISD---YI-NMWIFVTITNNRLNSRIYINGRLID 166
QY 203 SAETTGGAIREDDNNTILKDRCNNNNOYVSIKRFIEFCALNKEIEKLYTSLTFL 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 OKPTISNLGNHASNNINFEKDCDTHRYIWKIFENLEKELNKEIKLDYDQNSGIL 226
QY 263 RDPFGNPLRYDEYELLIPVASSKDYQKLN--ITDYATLNAPSYTNGKLNYYR-RLYN 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 KDFMGDLYQYKPYMLNLDYPNKYVDVNNNGIRGYWL-KGPRGSVWTTNLYNSSLYR 285
QY 320 GLKFLIRYPNNIDSPKSGDFIKLYSVNNNEHIYGVKDKGNAPFNNDRIILRVGNA 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 GTFKFLIKKYGASGNK-DNIVRNDRYVINVYKKEY-----RLATNA 326
QY 380 PGILYKMEAVKLRDLK-----TYSVOLKLYDDKNASIGLVGTHNGOI 423
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 SQAGVEKILSLALPVDGNLSQVYVMKSKNDGJTNNCKMNLQNNNGNDIGIFGH-QF 384

```

```

QY 424 GNDPNRDLILASNNYFNHLK--DKILGCDWYFVPTDEGM 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 NNIAK--LVASNNYNRQIERSSRTLCSWEIIPVDDGM 420

```

```

RESULT 3
US-08-981-087A-1
/ Sequence 1, Application US/08981087A
/ Patent No. US20020081304A1
/ GENERAL INFORMATION:
/ APPLICANT: Elmore, Michael J.
/ APPLICANT: Mauchline, Margaret L.
/ APPLICANT: Minton, Nigel P.
/ APPLICANT: Pasechnik, Vladimir A.
/ APPLICANT: Tibball, Richard W.
/ TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHAYE P.C.
/ STREET: 1100 No. US20020081304A1th Glebe Rd. 8th floor
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22201-4741
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/981,087A
/ FILING DATE: 27-MAY-1998
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/GB96/01409
/ FILING DATE: 12-JUN-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9511909.5
/ FILING DATE: 12-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crawford, Arthur R.
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 124-688
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-816-4100
/ TELEFAX: 703-816-4100
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 431 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-981-087A-1

```

```

Query Match 22.5%; Score 553; DB 8; Length 431;
Best Local Similarity 30.0%; Pred. No. 1.8e-34;
Matches 138; Conservative 92; Mismatches 152; Indels 78; Gaps 18;

```

```

QY 29 LKSTIINDINNDIISDISGNSVITYPDAOLVPGINGKAIHLVNNESSEVIYKAMD 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 IKDSIIMDKRENNKFDISGYSINISGVDVYI-STNRQOFGIYSSKPSVINAQND 77
QY 89 IEYDNENFTVSWLVRPKVASASHLEQYGTNEYSTISSMKKHSLSISGWSVSLKGN 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 ILYNGRQNFISFWRIPK---YFNKVNUNNEYTIIDCIARNN-----SGMKISLANYK 129
QY 148 LIWTLKDSAGEVROIFTR-----DLDPKFNAVLANKWVFTITNDRLSSANLYINGVLMG 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 IIMTLQOTQAGNNQKLVFYTQMISID---YI-NKWIIFVTITNNRLGNSRIYINGRLID 184
QY 203 SAETTGGAIREDDNNTILKDRCNNNNOYVSIKRFIEFCALNKEIEKLYTSLTFL 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

D6 185 EKSIINLGDHIVSINILFKYIGC--NDTRRYGIRFKEVFDELTGTELEPIYLSDEPPSIL 243

QY 263 RDPGNPLRYDTEYELLPVASSMDV-QLNINDYMLTNAPSTNCKLNIYR-KLYNG 320

D6 244 KDFGNVLLIYKRRIYLLNLRTDSSITQNSFNLN---INQGRGVYKQPMIFSNTRYLTG 299

QY 321 LKFTIKR-YTPNNEIDSFVKSQDF-----IKLY--VSYNNEHIYGVPKDGNA 365

D6 300 VEILIRKSGTSDISNTDNFVRKNDLAYINYVDROVEYRLADISIAPEKIIKLIRKSN 359

QY 366 FNNIDRLI---RVCYNAPGPIPLTKKMAVYKLRDLKITYSVOLKLIYDDKNASLGLVGTNGO 422

D6 360 NNSLGGQITLVDSISNNC-----TMNQNNNGGNIIGLGFHSNN 397

QY 423 IGNDPNDILIASINWYNNHLKDKIL--GCDWYFVPPDEG 460

D6 398 -----LVASWYVNNIRKNTSSNGCWTSTSKEHGW 428

```

1      RESULT 4
2      US-08-981-087A-3
3      : Sequence 3, Application US/08981087A
4      : Patent No. US20020081304A1
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Elmore, Michael J.
9      : APPLICANT: Mauchline, Margaret L.
10     : APPLICANT: Manton, Nigel P.
11     : APPLICANT: Pasechnik, Vladimir A.
12     : APPLICANT: Titball, Richard W.
13     : TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF
14     :
15     : NUMBER OF SEQUENCES: 6
16     :
17     : CORRESPONDENCE ADDRESS:
18     :
19     : ADDRESSEE: NIXON & VANDERHYE P.C.
20     : STREET: 1100 No. US20020081304A1th Giebe Rd. 8th floor
21     : City: Arlington
22     : STATE: VA
23     :
24     : COUNTRY: USA
25     :
26     : ZIP: 22201-4741
27     :
28     : COMPUTER READABLE FORM:
29     :
30     : MEDIUM TYPE: floppy disk
31     : COMPUTER: IBM PC compatible
32     : OPERATING SYSTEM: PC-DOS/MS-DOS
33     : SOFTWARE: PatentIn Release #1.0, Version #1.30
34     :
35     : CURRENT APPLICATION DATA:
36     :
37     : APPLICATION NUMBER: US/08/981,087A
38     :
39     : FILING DATE: 27-MAY-1998
40     :
41     : CLASSIFICATION: 424
42     :
43     : PRIOR APPLICATION DATA:
44     :
45     : APPLICATION NUMBER: PCT/GB96/01409
46     :
47     : FILING DATE: 12-JUN-1996
48     :
49     : PRIOR APPLICATION DATA:
50     :
51     : APPLICATION NUMBER: GB 9511909.5
52     :
53     : FILING DATE: 12-JUN-1995
54     :
55     : ATTORNEY/AGENT INFORMATION:
56     :
57     : NAME: Crawford, Arthur R.
58     : REGISTRATION NUMBER: 25,327
59     : REFERENCE/DOCKET NUMBER: 124-688
60     :
61     : TELECOMMUNICATION INFORMATION:
62     :
63     : TELEPHONE: 703-816-4000
64     :
65     : TELEFAX: 703-816-4100
66     :
67     : INFORMATION FOR SEQ ID NO: 3:
68     :
69     : SEQUENCE CHARACTERISTICS:
70     :
71     : LENGTH: 144 amino acids
72     : TYPE: amino acid
73     :
74     : STRANDEDNESS:
75     :
76     : TOPOLOGY: linear
77     :
78     : MOLECULE TYPE: peptide
79     :
80     : US-08-981-087A-3

```

Query Match	9.88;	Score 240.5;	DB 8;	Length 144;
Best Local Similarity	42.08;	Pred No.1.5e-11;		
Matches 47;	Conservative 22;	Mismatches 42;	Indels 1;	Gaps 1;

```

OY 177 NKWEITITNDRSSANLYINGVLMGSEITGLAIRDENNITLTKLDRCNNNOVYSIDK 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 NKMIPTVTTNNRLGSRIRYINGNLDKXSISNMDIHSDMLKRYGC -NDTRYVGLRY 73

OY 237 FRICKALNPREIEIKYTSYLSITFLDFPMNPLRYDTETYLILFVASSKDY 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 FKVDTELGKRIEYTLISDEDPILKQFWSNYLYLNKRRYLLNLTLDTSI 125

```

```

1      RESULT 5
2      US-08-981-087A-2
3      ; Sequence 2, Application US/08981087A
4      ; Patent No. US20020081304A1
5      ;
6      GENERAL INFORMATION:
7      ;
8      APPLICANT: Elmore, Michael J.
9      APPLICANT: Mauchline, Margaret L.
10     APPLICANT: Minton, Nigel P.
11     APPLICANT: Pasechnik, Vladimir A.
12     APPLICANT: Titball, Richard W.
13     TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF
14     NUMBER OF SEQUENCES: 6
15     CORRESPONDENCE ADDRESS:
16     ADDRESSEE: NIXON & VANDERHUYE P.C.
17     STREET: 1100 No. US20020081304A1th Glebe Rd. 8th floor
18     CITY: Arlington
19     STATE: VA
20     COUNTRY: USA
21     ZIP: 22201-4741
22     COMPUTER READABLE FORM:
23     MEDIUM TYPE: Floppy disk
24     COMPUTER: IBM PC compatible
25     OPERATING SYSTEM: PC-DOS/MS-DOS
26     SOFTWARE: PatentIn Release #1.0, Version #1.30
27     CURRENT APPLICATION DATA:
28     APPLICATION NUMBER: US/08/981,087A
29     FILING DATE: 27-MAY-1998
30     CLASSIFICATION: 424
31     PRIOR APPLICATION DATA:
32     APPLICATION NUMBER: PCT/GB96/01409
33     FILING DATE: 12-JUN-1996
34     PRIOR APPLICATION DATA:
35     APPLICATION NUMBER: GB 9511909.5
36     FILING DATE: 12-JUN-1995
37     ATTORNEY/AGENT INFORMATION:
38     NAME: Crawford, Arthur R.
39     REGISTRATION NUMBER: 25,327
40     REFERENCE/DOCKET NUMBER: 124-668
41     TELECOMMUNICATION INFORMATION:
42     TELEPHONE: 703-816-4000
43     TELEFAX: 703-816-4100
44     INFORMATION FOR SEQ ID NO: 2:
45     SEQUENCE CHARACTERISTICS:
46     LENGTH: 144 amino acids
47     TYPE: amino acid
48     STRANDEDNESS:
49     TOPOLOGY: linear
50     MOLECULE TYPE: peptide
51     US-08-981-087A-2

```

Query March	8.4%	Score 206	DB 8	Length 144
Best Local Similarity	33.6%	Pred. NO.	5.8e-09	
Matches 44	Conservative 35	Mismatches 46	Indels 10	Gaps 4
OY	29	LKSTIILNDINDIISDISGFSNVITYPDAQLVPGINGKRIHLNVNNESEVIAHKAMD	88	
		: : : : : :	: : : : : :	
Db	19	IKDSIILMRRENNKFDISGFSNISTNGDYIY-STRNQFGIYSSKPSPEVIAQNND	77	
OY	89	IEYNDMFNFNFVSPFLARYKVSASHLEQY-TNEYSIISMKHSLISGSSVSILKGN	147	
		: : : : : :	: : : : : :	
Db	78	ITVGRVGNFESISFVRIKPK---YFNKVNLTNETTIIIDICINNN---SGKKISLNYNK	129	
OY	148	LITWILKDSAGEVROIT	162	
		: : : : : :	: : : : : :	

Db 130 IITWLODPTAGNNOKL 144

## RESULT 6

US-08-981-087A-4  
; Sequence 4, Application US/08981087A  
; Patent No. US20020081304A1

GENERAL INFORMATION:

APPLICANT: Elmore, Michael J.

APPLICANT: Mauchline, Margaret L.

APPLICANT: Minton, Nigel P.

APPLICANT: Pesechuk, Vladimir A.

APPLICANT: Tjebkall, Richard W.

TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: NIXON & VANDERHAYE P.C.

STREET: 1100 No. US20020081304A1th Glabe Rd. 8th floor

CITY: Arlington

STATE: VA

COUNTRY: USA

ZIP: 22201-4741

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/981,087A

FILING DATE: 27-MAY-1998

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB96/01409

FILING DATE: 12-JUN-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9511909.5

FILING DATE: 12-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Crawford, Arthur R.

REGISTRATION NUMBER: 25,327

REFERENCE/DOCKET NUMBER: 124-688

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4100

TELEFAX: 703-816-4100

INFORMATION FOR SEQ. ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 143 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-981-087A-4

Query Match 5.6%; Score 137; DB 8; Length 143;  
Best Local Similarity 23.4%; Pred. No. 0.00089;  
Matches 40; Conservative 32; Mismatches 47; Indels 52; Gaps 8;

QY 311 NIYYR-RLVYGLKFLIR--YTPNNEIDSFYKSGD-----TKLV--VSYNNE 354  
Db 1 NITSNRLTYGVEYTIKKNGSTDISNTDFYRKNDLAYINVDQVEYKLADISIAKE 60  
QY 355 HAYGPRKGNANFNLDRL--RVGYNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNA 411  
Db 61 KIILKLTSSNNSLGLIIVDSIGNNC-----TMFNQNNNG 98  
QY 412 SLGLVGHNGOIGDPRNDILIASNMVFNLKDKIL--GCDWYFYPNDEGW 460  
Db 99 NIGLGHSMN-----LVASSWYNNIRKNTSSNGCFWSPISKHGW 140

RESULT 7  
US-09-797-097-2  
; Sequence 2, Application US/09797097

; Patent No. US20020058312A1  
; GENERAL INFORMATION:  
; APPLICANT: All, Benjamin  
; APPLICANT: Howard, John  
; TITLE OF INVENTION: Expression Cassettes and Methods of Delivery of Animal Vaccine  
; FILE REFERENCE: P002460U1  
; CURRENT APPLICATION NUMBER: US/09/797,097  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 08/529,006  
; PRIOR FILING DATE: 1995-09-15  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 2  
; LENGTH: 1447  
; TYPE: PRT  
; ORGANISM: Transmissible Gastroenteritis Virus Spike protein  
; US-09-797-097-2

Query Match 5.0%; Score 124; DB 10; Length 1447;  
Best Local Similarity 16.9%; Pred. No. 0.18;  
Matches 89; Conservative 76; Mismatches 162; Indels 200; Gaps 21;

QY 30 KSTILNDINDIDISIGFSSVITPDAGLVINGKALHLVNNESSEVIVHKAMD 89  
Db 354 KGAIVFSLNTTGGVLEISCIYNDIV-----SSSFSSYGEKPS 391  
QY 90 EYND-----MFNFTVSFWLRV---KVSASHLEQYGTNEYSISSMKKHSLSIGS 138  
Db 392 GVTDSRRCYVLYNCTALKYLTLPYKEIASKMGHYINGYFNFPIDCIS---- 447  
QY 139 WSVSLKGNLITWTLKDSAGEVROITFRDLPKFNAYLAKWFIITN---DLSSANLY 195  
Db 448 FNLTTGSDSDVEWTI-----ATSYTEALVQVENAITKVTYCNSTY 487  
QY 196 INGVL-----MGSAEI-----TGCAIRED----- 215  
Db 488 VNNIKCSQTLANLNGFVYSSEVGLVYKSVYLLPFTHTITVNTTIGLGMRSVGP 547  
QY 216 ---NNTLKDRCNQNNQVYSIDKFRICKALNPKIEIKLYSYLSITFLRDEWGNPL 270  
Db 548 IASTLSNLTLPQDNNNTDVCIRSDQFSV-----YVHSTCTSLMDNVF 591  
QY 271 RYDTEYLIYVASSSKQVQVKNITDYM-----YLTNAPSTNGKLNYYRLYN----- 319  
Db 592 K-----RNCSTDVLAATVAVIKTGTCPSEPKLNNYL--TNKKFCLSL 630  
QY 320 ---GLKFLIKRTPNNEIDSFYKSGDFIKLYVSYNNEHIVYPRKGNANFNLDRLR 374  
Db 631 SPVGNCKRFDAARTFN--DQVYRS-----LYVIEEGDNIVGVPDNGSLDLS- VLH 682  
QY 375 VG-----YVNPAGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASLGLV----- 416  
Db 683 IDSCTDYNITGRGVGIIRQTNRTLLSGYVTSGLDGLGFKVNSDVIVSYVPCDVSAQ 742  
QY 417 -----GTHNGQGNPNRNDILIASNM-----YFNHLKDKILG 448  
Db 743 AAVIDGTVGAI-TSINSELLGLTHWTTTPNPFYYSIYNTNDRTGR 788

RESULT 8  
US-09-801-574-57  
; Sequence 57, Application US/09801574  
; Patent No. US20020081592A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Peljing Jeremy  
; APPLICANT: Page, David C.  
; TITLE OF INVENTION: Reproduction-Specific Genes  
; FILE REFERENCE: 0399.2007-002  
; CURRENT APPLICATION NUMBER: US/09/801,574  
; CURRENT FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: 60/187,518  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: 60/261,557



; PRIOR FILING DATE: 2001-01-12  
 ; NUMBER OF SEQ ID NOS: 90  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 57  
 ; LENGTH: 2789  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-801-574-57

Query Match 4.7%; Score 116.5; DB 10; Length 2789;  
 Best Local Similarity 22.2%; Pred. No. 1.6;  
 Matches 111; Conservative 79; Mismatches 164; Indels 147; Gaps 30;

```

QY 3 FSPPIPSYKNDLDCWYDNEEDIVILKSTILMLDINNDIISDISGNSVITYPDAQ- 61
DB 1806 FSLSVPTGCVNFG---DSLEDFE-ILRKSTLK-----LINVCGSPVHSYFGKQD 1853
QY 62 ----LVGINKGKALHLYNNSSSEY-----IVHKAMDIEYMDNPNFVSR----- 102
DB 1854 HWMIIEMISSKVFNNNEAVRKISLYGLEHLFEDAKNLVWKERTQSFSKYSQKRD 1913
QY 103 ---WLRVPKVSASHLEQ-----YGTNEYSISSMKK-HSLSIGSGWGS 142
DB 1914 EERLLRYNKCASFSLKITYDLSKDLNNEPIISGLEDITIASRSDHPINEA---TIS 1970
QY 143 LK---GNL-----ITLKDLSA-----GEVROITFR-----DLPDKF----- 171
DB 1971 IENSKRNSNLAHPDICCISEILDQAEFADLKLQDLTLRCTDHLLEILKKYFQMLQDNM 2030
QY 172 -NAVLANKWFITITND-----RUSANLYINGVLMGSAEITGLAIRDNNTITLKD 223
DB 2031 DNFIIEENVLDVIVINSHAIILKPEALIMYLE-IWVSTTHFL-----KNSIARKLD 2084
QY 224 RCNNNNQVYSIDKFR--IFCKALNPKIEIKLYSYLSITFLRD-----FWGNPLRYDT 274
DB 2085 K-----QFRGMLWMDLSLP-ELVQCEKMASFELKNSYDVCIM----- 2125
QY 275 EYLLIPVASSSKVOUKNTDYMVLTINAPSYTGKLNITYRRL--YNGLEFIKRYPPNN 332
DB 2126 --KVITTAVSELKDDIDITICKINEAVNC-SYA--IHLSRELOELSEIKKLKK--SKY 2177
QY 333 EIDSPFKSGDFTKLIVSYNNNEHIVGYPKDNAPNNLRLILRGYNAPGPIYKMEAVK 392
DB 2178 FISTYI--DFVPYIASINGSVTELEYNYNOFSTL--LKNVMSAPRDLGKMAHIRK 2231
QY 393 LMDIKTYSQLALYDKNASL 413
DB 2232 V--MKTIE-HMKMICTKNAEL 2249
  
```

## RESULT 9

US-09-848-834A-14  
 ; Sequence 14, Application US/09848834A  
 ; Patent No. US20020076416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aphron Corporation  
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens  
 ; FILE REFERENCE: 1102865-0047  
 ; CURRENT APPLICATION NUMBER: US/09/848,834A  
 ; PRIOR FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: 60/202,328  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 14  
 ; LENGTH: 37  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G  
 ; OTHER INFORMATION: nrh hormone linked by a spacer to amino acid sequence 947-967 of  
 ; OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)  
 ; NAME/KEY: MOD\_RES

; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(10)  
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (11)..(16)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (17)..(37)  
 ; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor  
 ; OTHER INFORMATION: (Tentoxylisin)  
 US-09-848-834A-14

Query Match 4.6%; Score 114; DB 10; Length 37;  
 Best Local Similarity 95.5%; Pred. No. 0.0081;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 94 MFNNFTVSFWLRVPKVSASHLE 115
DB 16 LFNFTVSFWLRVPKVSASHLE 37
  
```

## RESULT 10

US-09-848-834A-18  
 ; Sequence 18, Application US/09848834A  
 ; Patent No. US20020076416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aphron Corporation  
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens  
 ; FILE REFERENCE: 1102865-0047  
 ; CURRENT APPLICATION NUMBER: US/09/848,834A  
 ; PRIOR FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: 60/202,328  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 18  
 ; LENGTH: 50  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum  
 ; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the  
 ; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a space  
 ; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (1)..(11)  
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (50)..(50)  
 ; OTHER INFORMATION: Amidated glycine or glycylamide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(10)  
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (11)..(16)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (17)..(37)  
 ; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (T  
 ; OTHER INFORMATION: oxylysine  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (38)..(41)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (42)..(50)  
 ; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
 US-09-848-834A-18

Query Match 4.6%; Score 114; DB 10; Length 50;  
 Best Local Similarity 95.5%; Pred. No. 0.012;  
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 94 MNFTVSWLRVPKVSASHLE 115  
:|||||  
Db 16 FNNFTVSWLRVPKVSASHLE 37

## RESULT 11

US-09-943-548-3  
; Sequence 3, Application US/09943548  
; Patent No. US20020042364A1  
; GENERAL INFORMATION:  
; APPLICANT: Rittershaus, Charles W.  
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY  
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 08/432,483  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: PCT/US96/06147  
; PRIOR FILING DATE: 1996-05-01  
; PRIOR APPLICATION NUMBER: 08/945,289  
; PRIOR FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: helper T cell epitope of tetanus toxin  
US-09-943-548-3

Query Match 4.6%; Score 112; DB 10; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 FNNFTVSWLRVPKVSASHLE 115  
:|||||  
Db 1 FNNFTVSWLRVPKVSASHLE 21

## RESULT 12

US-09-848-834A-4  
; Sequence 4, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Tetanus bacillus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(21)  
; OTHER INFORMATION: Amino acid sequence 947-967 of Tetanus  
; OTHER INFORMATION: Toxoid Precursor (Tentoxylysin)  
US-09-848-834A-4

Query Match 4.6%; Score 112; DB 10; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 FNNFTVSWLRVPKVSASHLE 115  
:|||||

Db 1 FNNFTVSWLRVPKVSASHLE 21

## RESULT 13

US-09-848-834A-10  
; Sequence 10, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 34  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of  
; OTHER INFORMATION: Tetanus toxoid precursor (tentoxylysin) linked by a spacer to  
; OTHER INFORMATION: Ino acid sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated phenylalanine  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(21)  
; OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor  
; OTHER INFORMATION: (tentoxylysin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (22)..(25)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (26)..(34)  
; OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone  
; NAME/KEY: MOD\_RES  
; LOCATION: (34)..(34)  
; OTHER INFORMATION: Amidated glycine or glycinamide  
US-09-848-834A-10

Query Match 4.6%; Score 112; DB 10; Length 34;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 FNNFTVSWLRVPKVSASHLE 115  
:|||||  
Db 1 FNNFTVSWLRVPKVSASHLE 21

## RESULT 14

US-09-815-242-12209  
; Sequence 12209, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: FLITRA 011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848

RESULT 15  
US-09-972-484-52  
Sequence 52, Application US/09972484  
Patent No. US20020127245A1  
GENERAL INFORMATION:  
APPLICANT: Miller, Timothy J.  
Kiepfer, Sharon  
Reed, Albert Paul  
Jones, Elaine V.  
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses  
Therefor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:

	Query Match	4.4%;	Score 108.5;	DB 10;	Length 1101;	
	Best Local Similarity	18.1%;	Pred. No. 1.8;	Mismatches 145;	Indels 155;	Gaps 21.
	Matches	83;	Conservative	75;		
QY	32	STILMDINDNDIISDISEGNSSVITPPDAOLVPGINGRAIHLLNNESSEFVIVHKMADIEY	91			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	8	ATVESLNTTGGVILEITESCYNDFV				SSSSFYSGEISF 43
QY	92	N-----DMFNFFVSEFLRV-----KVSASHLEOYGTNEYSITISMKKHSLSI---	135			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	44	GVTGDPFRCYALVNGTAKYLGTLPPSVKEIAISKMGHFYINGVNFSTFPIDCSFNLT	103			
QY	136	---GSGMSVSLAKG--NNLIWTLKDSAGEVRDITF-----RDLDPDK-----	171			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	104	TGDSGAFWTIAVSYTALVQVENTMA--IKVYTCNSHINNICKSQLTAFANLQNGYPPVAS	161			
QY	172	-NATLAKKWFVITTTNDRLSANLYINGVLMSGASETTLGAIRED-----NNTITLK	221			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	162	SEVGLVAKRSVYLLPSPFYSHTSVNITID-----LGMKRSGOYGPPIASTLSNITLP	210			
QY	222	LDRCNNNNQVYSIDKFEFLCKALNPKFELEKLYTSYLSITFLPDRDVGNDLPRLDYTEYLLIPV	281			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	211	MQ--DNNITD-----VYCIKRSNQ-----FSYVIVHSTCKSSIM-----	239			
QY	282	ASSSKDVOLKNITDYMVL-----TNAESYTGKLNITYRRLYN-----GLKRFII	325			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	240	----DVEFNSDCDGVLYATAFAVIKGTGCPFSFDKIANLT--TFNFCLSLANVGANCKEDV	293			
QY	326	KRYTPNNEIDSPFKSGGFILYYSNNNEHIVGYPKDGNFNNLDRIIRVAG-----YN	378			
	:	:	:	:	:	:
	:	:	:	:	:	:
Db	294	AARRTINE--QVVR-----LVYIEERGDNIVAPSPNSGHLHDS--VLHIDSCDTPNYNG	345			

Tue Dec 3 12:26:17 2002

us-09-816-467-2.rapb

Page 8

Oy 379 APGIPLYKKMEAVKLDLJKTYSVOLKLYDDKNASLGLV 416  
| : : : : | : : : : | : : : :  
Db 346 RTGVGIIIRQTNSTLLSGLYITSLSGDLGFKNVSDGVI 383

Search completed: December 1, 2002, 11:32:07  
Job time : 24 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:20:07 ; Search time 43 Seconds

(without alignments)  
1035.122 Million cell updates/sec

Title: US-09-816-467-2

Sequence: 1 MWFSTPFPSTSKNLDQWD.....DKILGDMYFVPTEGWTND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2451	99.8	1315 1	BRCLTN
2	653.5	26.6	1296 1	BRCLTN
3	644.5	26.2	1297 1	BRCLTN
4	643.5	26.2	1297 1	BRCLTN
5	617	25.1	1291 2	BRCLTN
6	590	24.0	1291 2	BRCLTN
7	578.5	23.6	1252 2	BRCLTN
8	571	23.2	1252 2	BRCLTN
9	560.5	22.8	1251 2	BRCLTN
10	517	21.1	1274 2	BRCLTN
11	430	17.5	1291 2	BRCLTN
12	430	17.5	1291 2	BRCLTN
13	411.5	16.8	1276 2	BRCLTN
14	406	16.5	1285 2	BRCLTN
15	269.5	11.0	1162 2	BRCLTN
16	263	10.7	1162 2	BRCLTN
17	244.5	10.0	1136 2	BRCLTN
18	244.5	10.0	1136 2	BRCLTN
19	244.5	10.0	1136 2	BRCLTN
20	217.5	8.9	1193 2	BRCLTN
21	217	8.8	1193 2	BRCLTN
22	214	8.7	1165 2	BRCLTN
23	212	8.6	1165 2	BRCLTN
24	201.5	8.2	1165 2	BRCLTN
25	143	5.8	369 2	BRCLTN
26	142	5.8	369 2	BRCLTN
27	136.5	5.6	4981 2	BRCLTN
28	134.5	5.5	1225 1	BRCLTN
29	134.5	5.5	1225 1	BRCLTN

30	132	5.4	1127	2	T8317	ORF MSV156 hypot
31	131.5	5.4	2925	2	T00133	RNA-directed RNA p
32	131	5.3	755	2	T41912	structural phospho
33	129	5.3	937	2	C97168	glycosyltransferas
34	129	5.3	1417	2	T18418	hypothetical prote
35	129	5.3	4436	2	E71086	hypothetical prote
36	128.5	5.2	1272	2	H82926	conserved hypotet
37	128	5.2	669	2	E71610	hypothetical prote
38	128	5.2	987	2	A64474	hypothetical prote
39	128	5.2	1225	2	A36607	hypothetical prote
40	128	5.2	1411	2	T18417	E2 glycoprotein
41	128	5.2	2523	2	T18417	hypothetical prote
42	127	5.2	1645	2	F98907	phage-related prot
43	127	5.2	2391	2	T18410	cardamoyl-phosphat
44	127	5.2	2496	2	A71616	secreted protein p
45	126.5	5.2	1132	2	H82887	hypothetical prote

## ALIGNMENTS

### RESULT 1

BRCLTN  
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin  
C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence-revision 31-Mar-1988 #text-change 03-Jun-2002

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

EMBL J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; MUID:87055814; PMID:3536478

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940770

A:Fairweather, N.F.; Lyness, V.A. Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747; PMID:3774547

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:940774

A:Fairweather, N.F.; Lyness, V.A. Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C

A:Reference number: A25194; MUID:86085672; PMID:3510187

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:9144920; PIDN:AAA23282.1; PID:9144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 461-475 <MAT>

A:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

A:Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: J50096; MUID:89093918; PMID:2463305

A:Contents: annotation; epitope region

R:Schlauer, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de lauroto, P.P.; Dasgupta,

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293; PMID:1331807

A:Contents: annotation

R:de Filipis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.  
Eur. J. Biochem. 229, 61-69, 1995  
A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.  
A:Reference number: S69348; MUID:95262688; PMID:7744050  
A:Accession: S69348  
A:Molecule type: protein  
A:Residues: 2-31 <DEF>  
C:Comment: The source of this protein was an extrachromosomal plasmid.  
C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fragment A) and heavy (fragment B) chains. The amino end of the heavy chain (fragment B) binds to ganglionic dual chains are not toxic when separated. Fragment C binds to peripheral neuronal synapses, is internalized and cleaved by proteolytic cleavage of synaptic vesicles.  
C:Function: This potent neurotoxin binds to peripheral neuronal synapses, is internalized and cleaved by proteolytic cleavage of synaptic vesicles.  
A:Description: blocks neuroexocytosis via hydrolysis of a Glu-Phe peptide bond in synaptobrevin.  
C:Superfamily: tetanus toxin  
C:Keywords: hydrolyase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
F:2-457/Product: tetanolysin light chain (fragment A) #status predicted <TL>  
F:461-1315/Product: tetanolysin heavy chain (fragment B) #status experimental <TX>  
F:461-864/DNA: channel forming (fragment B) #status predicted <TX>  
F:865-1315/DNA: channel forming (fragment C) #status predicted <TX>  
F:233-237/Binding site: zinc (His) #status predicted  
F:234/Active site: Glu #status predicted

Query Match 99.8%; Score 2451; DB 1; Length 1315;  
Best Local Similarity 100.0%; Pred. No. 6.3e-145; Indels 0; Gaps 0;  
Matches 462; Conservative 0; Mismatches 0;

QY 2 VSTPIPFYSKNDQVNEEDIVILKSTILINDINNDIISDGSFNSVITPPDAQ 61  
|||||  
DB 854 VSTPIPFYSKNDQVNEEDIVILKSTILINDINNDIISDGSFNSVITPPDAQ 913  
|||||  
QY 62 LVPNGKAIHLVNNSESEYVHKADIEYDNFNNFTYFVRLRVKVSASHLEQYGTNE 121  
|||||  
DB 914 LVPNGKAIHLVNNSESEYVHKADIEYDNFNNFTYFVRLRVKVSASHLEQYGTNE 973  
|||||  
QY 122 YSTISSMKHSLISIGGWSYSLKGNLWTLKDSAGEVROITFRDLPDFNFATLAKKWF 181  
|||||  
DB 974 YSTISSMKHSLISIGGWSYSLKGNLWTLKDSAGEVROITFRDLPDFNFATLAKKWF 1033  
|||||  
QY 182 ITTNDRLSSANLYINGVLMGSAETGLGAIREDNNITLKLDRCNNOYVSDKRIIFC 241  
|||||  
DB 1034 ITTNDRLSSANLYINGVLMGSAETGLGAIREDNNITLKLDRCNNOYVSDKRIIFC 1093  
|||||  
QY 242 KALNPKREIKLYSYSLTFLRDFMGNPLRDEYLLIPVASSSKDYLKNTIDYMLTN 301  
|||||  
DB 1094 KALNPKREIKLYSYSLTFLRDFMGNPLRDEYLLIPVASSSKDYLKNTIDYMLTN 1153  
|||||  
QY 302 APSYTNKLTIVYRRLNGLKFIIKRTYPNNEIDSEYKSGDFIKLYSVYNNNEHIVGPK 361  
|||||  
DB 1154 APSYTNKLTIVYRRLNGLKFIIKRTYPNNEIDSEYKSGDFIKLYSVYNNNEHIVGPK 1213  
|||||  
QY 362 DGAENNLDRILRYGVNAPGIPLYKKMEAVKLRDLKTYVQKLKYDKNASLGLVGHNG 421  
|||||  
DB 1214 DGAENNLDRILRYGVNAPGIPLYKKMEAVKLRDLKTYVQKLKYDKNASLGLVGHNG 1273  
|||||  
QY 422 QIGNDPMDILIASNMWFNLKDKILGCDMVFYPTDEGWTND 463  
|||||  
DB 1274 QIGNDPMDILIASNMWFNLKDKILGCDMVFYPTDEGWTND 1315  
|||||

RESULT 2  
BTCLAB  
bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum  
N:Alternate names: botulinum neurotoxin type A  
C:Species: Clostridium botulinum  
C:Date: 31-Mar-1993 #sequence, revision 31-Mar-1993 #text, change 18-Jun-1999  
C:Accession: A35294; S68820; A33401; A53884; A60025; A27000  
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.  
J. Biol. Chem. 265, 9153-9158, 1990  
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other  
A:Reference number: A35294; MUID:90264400; PMID:2160960  
A:Accession: A35294

A:Molecule type: DNA  
A:Residues: 1-1296 <IN>  
A:Cross-references: GB:M30196; NID:q144864; PIDN:AAA23262.1; PID:q144865  
A:Experimental source: Strain 62A, subtype A  
R:Thompson, D.E.; Brehm, J.R.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, R.; Thompson, D.E.; Brehm, J.R.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, R.  
Eur. J. Biochem. 189, 73-81, 1990  
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxin.  
A:Reference number: S09492; MUID:90235864; PMID:2185020  
A:Accession: S09492  
A:Molecule type: DNA  
A:Residues: 1, 'Q', 'V', '28-1296 <THO>  
A:Cross-references: EMBL:X52066; NID:q40381; PIDN:CAA36289.1; PID:q40382  
A:Experimental source: NCTC 2916  
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
FEBS Lett. 376, 41-44, 1995  
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compo  
A:Reference number: S67988; MUID:96096783; PMID:8521962  
A:Accession: S67988  
A:Molecule type: DNA  
A:Status: preliminary  
A:Residues: 1-12 <PUJ>  
A:Cross-references: EMBL:D67030; DBJ:D50421; NID:q2160224  
R:Belley, M.C.; Somers, E.; Dasgupta, B.R.  
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989  
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t  
A:Reference number: A33401; MUID:89350959; PMID:2669749  
A:Accession: A33401  
A:Molecule type: DNA  
A:Residues: 1-35 <BE>  
A:Cross-references: GB:M27892; NID:q144880; PIDN:AAA23269.1; PID:q551776  
R:Giannini, J.A.; Dasgupta, B.R.  
J. Protein Chem. 12, 351-363, 1993  
A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,  
A:Reference number: A53884; MUID:94000342; PMID:8397793  
A:Accession: A53884  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 867-880;1148-1217, 'Y', 1219 <GIN>  
A:Experimental source: strain Hall  
R:Dasgupta, B.R.; Foley, J.; Niece, R.  
Biochemistry 26, 4162, 1987  
A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.  
A:Reference number: A27000  
A:Accession: A27000  
A:Molecule type: protein  
A:Residues: 2-47 <DAS2>  
R:Binz, T.; Blas, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.  
J. Biol. Chem. 269, 1617-1620, 1994  
A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.  
A:Reference number: A49708; MUID:94124495; PMID:8294407  
A:Contents: annotation  
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy  
C:Genetics:  
A:Gene: atx; botA  
C:Function: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associ  
C:Superfamily: tetanus toxin  
C:Keywords: disulfide bond; hydrolyase; metalloproteinase; neurotoxin; transmembrane p  
F:2-444/Product: bontoxilysin A light chain #status experimental <LGH>  
F:445-1296/Product: bontoxilysin A heavy chain #status experimental <HNY>  
F:223,227/Binding site: zinc (His) #status predicted  
F:224/Active site: Glu #status predicted

Query Match 26.6%; Score 653.5; DB 1; Length 1296;  
Best Local Similarity 33.1%; Pred. No. 8.6e-33;



Db 1183 NVVKNKEY-----RLATNASOAGVEKILSALEIPDVGNLSQVVMK 1224  
 QY 401 -----VOLKLYDDKNASLGLVGFTHNGCIGNDPDRDILASNMWFNHL--KDKILG 448  
 Db 1225 SKDDGIRNKKMKNLQDNGNDIGFIFGLHYD-----NLAFLVASNMWNRQVGNASRFEG 1279  
 QY 449 CDWYFVPTDEGM 460  
 Db 1280 CSWEFIPVDGGM 1291

## RESULT 5

140631  
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 12-Aug-1996 #sequence:revision 12-Aug-1996 #text\_change 16-Jul-1999  
 C:Accession: I40631; S48103; S48104; S36015  
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
 A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum  
 A:Reference number: I40631; MUID:94122659; PMID:7764370  
 A:Accession: I40631  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <RES>  
 A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J: Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id  
 A:Reference number: S48103; MUID:94013372; PMID:8408542  
 A:Accession: S48103  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 634-761, 'E', '763-841, 'M', '843, 'T', '845, 'N', '847-994 <CAM1>  
 A:Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779  
 A:Experimental source: non-proteolytic strain 2129B (Scott)  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
 A:Accession: S48104  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-843, 'T', '845, 'N', '847-994 <CAM2>  
 A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781  
 A:Experimental source: non-proteolytic strain Ekund 2B (Colworth 229)  
 C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic S  
 C:Genetics:  
 A:Gene: bontC/b  
 C:Superfamily: tetanus toxin  
 C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc  
 F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>  
 F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>  
 F:230-234/Binding site: zinc (His) #status predicted  
 F:231/Active site: Glu #status predicted

Query Match 25.1%; Score 617; DB 2; Length 1291;

Best Local Similarity 31.0%; Pred. No. 1.6e-30;  
 Matches 149; Conservative 94; Mismatches 191; Indels 46; Gaps 14;

QY 5 TP1PEFSYKSLDQWDEEDIDVILK-----KSTLINDINDISDISGENSESV 54  
 Db 834 TL1PEFLS-----TYSIEILIKFKNYSEILNLLNLLRNNLLIDLSGYAKV 885  
 QY 55 ITPYDPAOLVGINCK-AIHLVNNESSEVIYKAMDEIYNDMFNNFVSEMLRPKVSASH 113  
 Db 886 EYDVGVL-----NDKNQFKLTSASDKIRVTONIIFNSMFLEDFSVSFWRIRPKYRND 941  
 QY 114 LEQGTREYSITSMKHHSLISGSGSVSLKGNLLIWLKDSAGEVQJFR-DLPKFN 172  
 Db 942 IQNYIHREYITLNCMKNN-----SGMKTISRGKRIITWLIDINKTYSVFEEYIREDIS 996  
 QY 173 AYLANKVVFTITNDRLSSANLYINGVLSAETTGATREDNNITLKDRCNNNOYV 232  
 Db 997 EYI-NRNFVYITNN-LDNATIIYNGTLESNMDIKDIGEIVNCEILFKFDGDVDRQFI 1054

QY 233 SIDKRIEFCALNPKEIEKLTYSITSLIFELRDFWGNPLRDTVEYLLIPVASSKDVOLKN 292  
 Db 1055 WKYFSTFNTOLNOSNKEIKYIOSYSEYLDKDFNCPMLMYKKEYMFAAGKNKSYIKLVK 1114  
 QY 293 ITDYMVLTNAPSYTNGKLNITYYRRLYNGLKFIKRYTPNNEI-DSFVKSDFIKLYSVN 351  
 Db 1115 DSSVGEILIRSKYQNSNYINRYNLYICEKFLIRRESQSQINDIVREKEDIHLDLVH 1174  
 QY 352 NNE-HIVGPRKDNAPFNLDRLIRGVNAPGIPLYKKMEAVKLRDLKTYSVOLKLYDDKN 410  
 Db 1175 HEEMRVYAV---KYFKDDEKFLSTISDSNEFYKTIKRYEDQPSYSCOLLKDEE 1230  
 QY 411 AS--LGLVGFTH---NGCIGNDPDRDILIASNMWFNHLKDK---ILGDMYFVPTDEGMT 461  
 Db 1231 STDIGLIGIRHFYSGLVRKKYKDYFCISKWYKLEVKRRPKYSNLGCMWCFIPRDEGMT 1290

## RESULT 6

A48940  
 bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum  
 N:Alternate names: botulinum neurotoxin type B (BoNT/B)  
 C:Species: Clostridium botulinum  
 C:Date: 19-Dec-1993 #sequence:revision 18-Nov-1994 #text\_change 18-Jun-1999  
 C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574  
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Breilm, J.K.; Atkinson, T.; Minton, N.P  
 Appl. Environ. Microbiol. 58, 2345-2354, 1992  
 A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the  
 A:Reference number: A48940; MUID:92384550; PMID:1514783  
 A:Accession: A48940  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <WHB>  
 A:Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735  
 A:Experimental source: type B, Danish  
 A:Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBI:P:112081); this publ  
 R:Campbell, K.D.; Collins, J.D.; East, A.K.  
 J: Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific  
 A:Reference number: S48103; MUID:94013372; PMID:8408542  
 A:Accession: S48105  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-994 <CAM>  
 A:Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783  
 A:Experimental source: proteolytic type B, strain NCTC 7273  
 R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.  
 submitted to the EMBL Data Library, April 1992  
 A:Description: Partial amino acid sequence of botulinum neurotoxin type B and compari  
 A:Reference number: S21575  
 A:Accession: S21575  
 A:Molecule type: DNA  
 A:Residues: 36-217, 'G', '219-224, 'S', '226-246 <SZB>  
 A:Cross-references: EMBL:211934; NID:g40383; PIDN:CAA77991.1; PID:g40384  
 R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars  
 J. Biol. Chem. 267, 14721-14729, 1992  
 A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus  
 A:Reference number: A42871; MUID:92340509; PMID:1634516  
 A:Accession: A42871  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-313, 'S', '315-451 <KUR>  
 A:Experimental source: strain Okra  
 A:Note: sequence extracted from NCBI backbone (NCBI:P:109365)  
 R:Dasgupta, B.R.; Datta, A.  
 Biochimie 70, 811-817, 1988  
 A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity w  
 A:Reference number: S07155; MUID:89000987; PMID:3139097  
 A:Accession: S07155  
 A:Molecule type: protein  
 A:Residues: 2-29, 'M', '31-45 <DAS>  
 A:Accession: S08562  
 A:Molecule type: protein  
 A:Residues: 442-463, 'R', '465-467 <DA2>



	Query Match	24.0%;	Score 590;	DB 1;	Length 1291;	
	Best Local Similarity	31.7%;	Pred. No. 7.8e-29;			
	Matches 140;	Conservative 89;	Mismatches 184;	Indels 28;	Gaps 12;	
QY	34 ILINDINNDIISDISGNSSVITYPPDPAOLVPGINGK-AHLVNNESSVEIVHKRADIENX 92					
Db	865 ILLNRYKDNNLIDLDSYGAKVEYYDDGL-----NDKNQFKLTSSANSKRITVQNONTITN 920	:	:	:	:	:
QY	93 DMFNFTYSFWLRVRPKVSASHLEQYGTNEYSIISMRRKHSLSIGSGWSVLKGNNLIWTL 152	:	:	:	:	:
Db	921 SVFLDESVSFEWIRIPRYKNDGIQNYIHNEYTIINCMMKN-----SGMKISRGRNIWTL 975	:	:	:	:	:
QY	153 KDSAGEVQIFER-DLPDKFNATLANRWPFITTTNDLSANLYINGVLMGSAETTGCA 211	:	:	:	:	:
Db	976 IDINKTGSVFEEFYRIREDISEYI-NRWEFVTITNN-LNNAKIYNGKLSENTDLIKDIRE 1033	:	:	:	:	:
QY	212 IREDNNITLKLDRCNNNNNOYVSDIKRFIECKALNPKEIEKIYTYSYLSTFLRDENGPNLR 271	:	:	:	:	:
Db	1034 VIANGELIFEKLDGDIDDRQFIWKMYTIFRTTELQSOSIEEYKIQSYSEYIKDFEGNPLM 1093	:	:	:	:	:
QY	272 YDTEYYILPVASSSKSDVOALKNITDYMLTNAPSYTNGKLNYYRLNGLTFFIIKRYTPN 331	:	:	:	:	:
Db	1094 YNKKEYMFNAGKNRSYIKLKKDSPVGEILLTRSKYNQNSKYINYNDLVIGEKFIIIRKKS 1153	:	:	:	:	:
QY	332 NEI-DSFYKSGFILT-YSYANNNEHVIQYPKDGMAFNNDRIIRVGYNADGPIPLYKME 389	:	:	:	:	:
Db	1154 QSIDNDVIRKREYIYILDFFNLQEMWRVTV---KYFKKEEEKLF LAPISDSDEFYNTIQ 1209	:	:	:	:	:
QY	390 AVKLRDELTKTSVOLKLYDDKNAS--LGIVGTH--NQIIGNDPARDDLILASNWFNHDKD 444	:	:	:	:	:
Db	1210 IKKEYDEQFTYSCQLLFKKADESTDEIGLIGHREFYESGVIFEYKEDYCISKWYLEKVR 1269	:	:	:	:	:
QY	445 K----ILGCDWYFVPPTDEGWT 461	:	:	:	:	:
Db	1270 KPYNLKLGCNMNQFIRPKDEGWT 1290	:	:	:	:	:
RESULT 7						
S21178	botulinum neurotoxin type E precursor - Clostridium botulinum					

	Query Match	23.6%;	Score 578.5;	DB 2;	Length 1252;
	Best local Similarity	30.1%;	Pred. No. 3.9e-28;		
	Matches 148;	Conservative 82;	Mismatches 159;	Indels 103;	Gaps 16;
QY	7	IPFSYKNDGCVNDEEDIDYL-----KSTIIINDINDIISDISGFNSSVI	55		
Db	823	IPFLLSYTD-----DKLLISYFNKFPRIKSSVLMNRYKKNKYVDTSQDSHIN	873		
QY	56	TYPAQLVPGINGKAIHLVNNESSSEVIYHKAMDIEVDNENNFTEVSPWLVPKXSATLE	115		
Db	874	INGDVYKYP-TNKKQFGIYNCKLSEVNITSQNDIITYDNKTKNFSISWVRIIPNDKNIYN	932		
QY	116	QYGTNESEIISSMKKHSLISIGSGWSVSLKGNLILWTLKDSAGEVROTFTR-DLPDKENAY	174		
Db	933	V--NNEYTIIINCMRDNN---SGMKVSLNHNEIITLQDNAGINQKLAIFYGNANGISDY	986		

[illegible]

RESULT 8  
S33411  
botulinum neurotoxin type F - Clostridium baratii  
C.Species: Clostridium baratii  
C.Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
C.Accession: S33411, S31860  
R.Thompson, D.E.; Hutson, R.A.; East, A.R.; Allaway, D.; Collins, M.D.; Richardson, P.T  
FEMS Microbiol. Lett. 108, 175-182, 1993  
A.Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin  
A.Reference number: S33411; MUID:93252228; PMID:8468245  
A.Accession: S33411  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-1268 <THO>  
A.Cross-references: EMBL:X68262; NID:949138; PIDN:CAA48329.1; PID:949139  
C.Superfamily: tetanus toxin  
C.Keywords: neurotoxin

```

Query Match      32.2%; Score 571; DB 2; Length 1268;
Best Local Similarity 3.1%; Pred. No. 1.2e-27;
Matches 153; Conservative 90; Mismatches 168; Indels 66; Gaps 19;

Qy      7 IPESYSKNLDOWNEDIEDILK-----KSTIINDINDIISDISENSSYIYPDA 60
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
Db      833 IPELSE-----YTNDKILIHILIREYKRRIIDSSILMKYENNREIFDSSGNSISINGDI 888

Qy      61 QLYPGINGKAIHLVNNESSEVIYHKAMDIENYDMENNFVSTFLRPVKYSASHEQGTN 120
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      889 YIY-STNRNQGFIYSRSLEEVNITQNTIIYNSRQNFSEVSEWVRIPKY--NNLKNL--NN 944

Qy      121 EYTISSMKKSHLSIGSSRSVSLKGNNLITWLTKDSAGEVROIFR-----DLDPKFAAYL 175
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
Db      945 EYTLINCMRRNN-----SGWKISLNNYNNIIMTLQDTGNOKKLFENYTMQIDISD----YI 996

Qy      176 ANKVFELTITDRLSSAALYINGVLMSAEITGLGAIREDNNITTLKDRCNNNQAYSID 235
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
Db      997 -NKWTFVTTITNNRIGHSHKLYINGNLTDOKSIINLGINIHVDNLIIFKIVGC-NDRIRYIGIR 1054

Qy      236 KFRFKCALNKEIEFKLTSYLSTIFELRFMGNPLELYDFREYELI---PAASSKDYOLK 291
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
Db      1055 YFKFNNMELDTLEIETLYHSEDPSTIILKDFMGNYLLNNKKYYILNLNLKPKMNSYVRKNSDIL 1114

Qy      292 NITDYMFLTNAPSYTNGKLNIIYR--RLYNGLFLFIIRY--TPANNEIDSYKSGDFIKLYV 348
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
Db      1115 NI-----NNQRCIYSTNIEPSNARLYGVEVILIRKVGSTDFGSPFVRKNDTYIVINV 1167

Qy      349 SYNNEHIVGYPKDGNAFN---NLDRIKRVGYNAPGIPLYKKMEAYVAKRLDKITYSVQLKL 405
      ||| | : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      1108  VDGANEYQLADYVSISAVEKTIKLRIRISMSNYNSQMIIINDSLGDNCTMNEKT-----1220
QY      406   YDKRASLGLVGTHTNQIGNDENRDILLIASNWFYHNLKDKIL--GCDWYEVPTDEGW 460
        :      :      :      :      :      :      :      :      :      :
Db      1221  --NNGNDIGLGFH--LNN-----LVASWYWKINRNTRNNGCGWFSLSKEHW 1266

```

## RESULT

botulinum neurotoxin type E precursor - Clostridium butyricum  
C:Species: Clostridium butyricum  
C:Date: 30-Jun-1992 #sequence\_revision 15-May-1998 #text\_change 16-Jul-1999  
C:Accession: JH0256; S16145  
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.  
Biochem. Biophys. Res. Commun. 183, 107-113, 1992  
C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release from the presynaptic terminal by blocking the binding of toxin to cell receptors while the C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin  
F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>  
F:1423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>  
F:412-426/Disulfide bonds: #status predicted

```

Query Match 12.8%; Score 560.5; Db 2; Length 1251;
Best Local Similarity 29.3%; Pred. No.5,26-27;
Matches 147; Conservative 83; Mismatches 164; Indels 107; Gaps 17;

QY      1 MVEST---PIPEFSYKMLDCWVNEDEDIVL-----KSTIILNDLINDIISD 46
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      814 MIDTLNNSIPFKLSSYDP-----DKLISYFNKFEFKRIKSSSVLMRYKNDXYVD 864

QY      47 ISGNSSVITTPDQOLVINGINKAHLVNNSESSEVIVHKAIDIEYNDMFNNFTVSPMLRV 106
      |||::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      865 TSGDSNININGDYKYP-TNRNDFGILNIDKLEVAISONDYIITDKRYKFNFSISFVRI 923

QY      107 PKVSASHLEQGYTVEYSITSSKKHSLSGSWSLKGNNLIWTLKDSAGEVROQTFR- 165
      |::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      924 PNYNKLIVN--NNEYTLINCRDNN---SGWKVSLNHEIITWTLQDNGINOKLAFNY 977

QY      166 DLPDKFNAYLANKVFTITTDRLSSANLYINGVLKGSMAETTLGAREDDNITLKDRC 225
      ::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      978 GNANGISDLY-NKWLFEVITTDRLDGLSKLYINGNLIDKSSILNGLNHHVSDNLIFFKLYNC 1036

QY      226 NNNQOYSIDKFRIFCKALPKAEIKLYTSYSLTFELDPFGMPLRYDTEYLLIPVAS-- 283
      ::||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      1037 -SYTRYIGIRFENIFDKLEDEIQTLYNNPRANLTKDPMGWYLLVDKDEYLLNLVLPKN 1095

QY      284 -----SKRDVQLNITDYMVLNAPSYTNGKLNLYRRLYNGIKELIKRTYPNNEIDSF 337
      |||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      1096 NFINRRDSTSLINNIRSTILLAN-----RLYSGIKVLIQRNNSSSTNML 1141

QY      338 VKSGDEFILVLSY-----NNNEHIVGYPKDGAAFNMLDRLRLRGYNAPG 381
      ||::||::||::||::||::||::||::||::||::||::||::||::||:
DB      1142 VKRND--GVYINFAVASKTHLLPLADYATATTNKKRTIKISSGNRFNV-----VYMSVVG 1194

```

OY 382 IPIKKMEAVKRLDQKTYTVOQLKLYDKNMSLGLVGHNGOIGDNDPNDILIASWYENH 441  
 Db 1195 -----NCTMNFKNMNNNGNIGLGF-----KADTVASTWYTH 1227  
 OY 442 LKDKI--LGDWYFVPTDEGW 460  
 Db 1228 MRODTNSNGFFWNISEBHW 1248

## RESULT 10

140813  
 neurotoxin type F - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Jul-1999  
 C:Accession: I40813; S48108  
 R:Past, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.E.  
 FEMS Microbiol. Lett. 96, 225-230, 1992  
 A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.  
 A:Reference number: I40644  
 A:Accession: I40813  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1274 <RES>  
 A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J. Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id  
 A:Reference number: S48103; MUID:94013372; PMID:8408542  
 A:Accession: S48108  
 A:Status: preliminary; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 634-1002 <CAM>  
 A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin

Query Match 21.1%; Score 517; DB 2; Length 1274;  
 Best Local Similarity 29.1%; Pred. No. 2.7e-24;  
 Matches 131; Conservative 99; Mismatches 160; Indels 60; Gaps 17;

OY 29 LKSTLNDINDIISDISEGNSVITPPDQVLDPQINGKA-----IHLVNESS 79  
 Db 865 IKDSSILDMRRENNKFIIDISYGSNI-----SINGNVYISTNRNOFGIYNRLS 914  
 OY 80 EYIVHAKMIEYNDMNNNTVSPFLRVPKYSASHLDQYGN EYIISMKKHSLSIGSG 138  
 Db 915 EYVIAQNDIITNSRIQNSISISFWRIPK---HYKPMNHNREYTIINCNGNNV---SG 966  
 OY 139 MSVSL---KGNMLIWTLKDSAGEVROITER-DLPKFNAYLANKWVFITITNRLSSANL 194  
 Db 967 WKISLRTVADCELIWTLQDSGKKEHLIRYEBELNLSNIT-NKWIFTVITTNRLGNSRI 1025  
 OY 195 YINGVLMSGAETTGATIEDNNITLKLDRONNNNOYVIDKFRIFCKALNPKREIEKLYT 254  
 Db 1026 YINGMNLIVKRSISNIGDHSVSNILFKIIVGCDDET-YVGIRYKVPFTEDEKTEIEFTLS 1084  
 OY 255 SYLSTFLTEDFGNGPLRKYTEYLLIPVASSSKDYOVLKNITDYMVLYLNAPSYTGKLNITY 314  
 Db 1085 NEPDPSILKNTYNGATILNKKYTLLENLKRDKYITLNL-SGIININQOQRYTEGSAFLNLY 1142  
 OY 315 RRLVGLKFLIIRKYP--NNEIDSEYKSGDFIKLYSYNNNEHIVGYPKDGNAFNNDRI 372  
 Db 1143 -KLYGVEVILIRKNGPIDISNDNFVRKNDL--AYI--NVVDKGVAYRLXADPKSEKEKI 1197  
 OY 373 LRVGNAGGIPLYKKMEAVKRLDQKTYTVOQLKLYDKNMSLGLVGHNGOIGDNDPNDIL 432  
 Db 1198 IRTS-----NLNDLSGQIIVWDSIGNNCTMNFONNNGSNIGLGFHSNN-----L 1242  
 OY 433 IASWYENHKLKDKIL--GCDWYFVPTDEGW 460  
 Db 1243 VASSWYNNIRNRNINSSNGCFMSSISKENGW 1272

## RESULT 11

S46431  
 botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)  
 M:Alternate names: BONT/C1 protein  
 C:Species: Clostridium botulinum phage 1C  
 A:Variatey: Strain C 468  
 C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 16-Jul-1999  
 C:Accession: S46431; S49107  
 R:Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.  
 Mol. Gen. Genet. 243, 631-640, 1994  
 A:Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxi  
 A:Reference number: S46426; MUID:94301293; PMID:8028579  
 A:Accession: S46431  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <HAN>  
 A:Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175  
 A:Experimental source: strain C 468  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993  
 C:Superfamily: tetanus toxin

Query Match 17.5%; Score 430; DB 2; Length 1291;  
 Best Local Similarity 26.4%; Pred. No. 7.5e-19;  
 Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

OY 3 FSTPIPFYSKNIIDCWVNDNEEDIDVI-----LKKSTLNDINDIISDISEGNSVIT 56  
 Db 840 FQNTIPF-----NIFSTNNSLKDIIINEFNNINDSKILSLONKRNKTLVPTSGINAEVSE 895  
 OY 57 YPDQVLV-----GINKAIIHLVNNESSEVIVHAKMIEYNDMNNNTVSPFLRVPK 108  
 Db 896 EBDVQVLPPIFPDFKLGSSGE-----DRGVIVQNEVIVNNSWYESPISFWINK 948  
 OY 109 VSASHLEQYGTNEYSITSSMKKHSLSIGSGWSYSLGNNMLIWTLKDSAGEVROITER-DL 167  
 Db 949 -WVSNLP-----GYTITDSVKNK---SGWSIGIISNLFVLFLKONEDSEOSINFSDYI 997  
 OY 168 PDKFNAYLANKWVFITITNDRLSANLYINGVLMGSAEITGLCAIFREDNNITLKLDR--- 224  
 Db 998 SNNAPGY--NKWFYVTVTNMNMGNMKIYINGKLIIDIRYKELGINSFKITTEINKIPD 1055  
 OY 225 -----CANNNOYVSIDKFRIFCKALNPKREIEKLYSYLSTFLRDFGNGPLRKYTEYLLI 279  
 Db 1056 TGLITSDSDNINMMIRDFYIFAKELDKDINILFNSLQYTNVYKDYWGNDLRYNKEYYV 1115  
 OY 280 PVASSSKDYOVLKNITDYMVLYLNAPSYTGKLNITYRLTN-----GLKFIIRKYPNNEI 334  
 Db 1116 NI-----DYL---NRMYANSRQIVFNTRRNNNDPNEGKTIIRKIRGNTN- 1158  
 OY 335 DSVKSGDFIKLYSYNNNEHIVGYPKDGNAFNNDRIILRVGNAGGIPLYKKMEAVK-- 392  
 Db 1159 DTRVAGDILYEDMTINNKAY-----NLPMKNETVADHNSHTEIDYALIGRQETKDI 1210  
 OY 393 -----LRDLKTYSVOL--KIYDKNMS-LGLVGHNGOIGDNDPNDILIASWY- 438  
 Db 1211 NDNIIFQIOPMNNNTYVASQLFKSNFNGENISGISIGYRRLGDD-----WVR 1260  
 OY 439 FNHLKDKI-----LGDWYFVPTDE 458  
 Db 1261 HNYLVPTVYKQGNVYASLSESTTHMGFPVSE 1291

## RESULT 12

A49777  
 botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain  
 C:Species: Clostridium botulinum phage  
 C:Date: 10-Mar-1994 #sequence\_revision 07-Apr-1994 #text\_change 23-Mar-2001  
 C:Accession: S11291; A35396; S22166; A49777  
 R:Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet,  
 Nucleic Acids Res. 18, 4924, 1990  
 A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.  
 A:Reference number: S11291; MUID:90370487; PMID:2204031  
 A:Accession: S11291

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-84, 'P', 86-1291 <HAU>  
 A:Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906  
 R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.  
 Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990  
 A:Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin  
 A:Reference number: A35396; MUID:9102498; PMID:222245  
 A:Accession: A35396  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-669, 'R', 671-1291 <TS1>  
 R:Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.  
 submitted to the EMBL Data Library, December 1991  
 A:Description: Nucleotide sequence of the gene for one of the components of hemagglutinin  
 A:Reference number: S22163  
 A:Accession: S22166  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <TS2>  
 A:Cross-references: EMBL:X62389; NID:g558175; PIDN:CAA44263.1; PID:g40390  
 R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.  
 Appl. Environ. Microbiol. 57, 1168-1172, 1991  
 A:Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and who  
 A:Reference number: A49777; MUID:91282468; PMID:2059039  
 A:Accession: A49777  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-607 <TS3>  
 A:Cross-references: GB:090210  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin

Query Match 17.5%; Score 430; DB 2; Length 1291;  
 Best Local Similarity 26.4%; Pred. No. 7.5e-19;  
 Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

OY 3 ESTPIPEFSYKNDGWDNEDIDVI-----LKKSTILNDINDNDIISDSCFNSSVIT 56  
 DB 840 FQMTIP-----NIFSYTNNSLKDIIIEYNFNNINDSKILSLONKKNLIVDTSGNAEYSE 895  
 OY 57 YPDQOLVP-----GINKKATHLVNNESEVIVHKAMDIENDMENFTVSPFLRVPK 108  
 DB 896 EGVOLNPIPFPPFKLGSSGE-----DRGKVIPTQENINIVNSMEFSISFWIRNK 948  
 OY 109 VSAASHLEQYGTNEYSTISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFR-DL 167  
 DB 949 -WVSNNLP-----GYTIIDSVKNN-----SGMSIGIISNFIPLVFLTKQEDSEOSINFSYDI 997  
 OY 168 PDKFNAYLANKWVEITITNRLSSANLYINGVLMSAEITGLAIREDNITLKLDR--- 224  
 DB 998 SNNAPGT--NKWFFVYITNNMGMKIKYINGKLIDITKVKELTGINSKTTPELTKRPD 1055  
 OY 225 -----CNNNNQVYSIDKFRICFALKNEIEKLYTSYSTIFELDFWGNPLRYDTEYLI 279  
 DB 1056 TGLITSDSDNINMMIRDFYIFAKELDKDINILFNSLOQYTVVAVKQNDLRYNKEYYAV 1115  
 OY 280 PVASSKSDVOLKNTDQWYLTNAPSYNGKINITYRRLYN-----GKFEIIRKRYPNNEI 334  
 DB 1116 NI-----DYL-----NRYMANSROIIVNTRRNNDPNEGKIIITIKRIGNTN- 1158  
 OY 335 DSFEKSGDFIKLYSVNNNEHIVGPRKGNFNNLDRILRVGNAPGIPLYKKMEAVR-- 392  
 DB 1159 DTRVRGDDILYFDMTINKKAV-----NLFMKNETPVYADNNHSTEDYIATGLRFQTKDI 1210  
 OY 393 -----LRDLKTYSVOL--KLYDDKNAAS-LGLVGTNAGQGNPNRDLILASNNY- 438  
 DB 1211 NDNIIFQIOPNNNTYVYASQIFKSNFNGENISGICSTIGTYRFLGGD-----WYR 1260  
 OY 439 FNHLKDKI-----LGCDDWVFVPTDE 458  
 DB 1261 HNYLVPTVYKQGNVYASLLESTSTHNGFVYVSE 1291

RESULT 13  
 S11455  
 botulinum neurotoxin type D - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
 A:Accession: S11455  
 R:Binz, T.; Kurazono, H.; Popoff, M.R.; Eklund, M.W.; Sakaguchi, G.; Kozaki, S.; Krie  
 Nucleic Acids Res. 18, 5556, 1990  
 A:Title: Nucleotide sequence of the gene encoding Clostridium botulinum neurotoxin ty  
 A:Reference number: S11455; MUID:91016853; PMID:2216736  
 A:Accession: S11455  
 A:Status: preliminary; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1276 <IN>  
 A:Cross-references: EMBL:X54254; NID:g40395; PIDN:CAA38175.1; PID:g40396  
 C:Superfamily: tetanus toxin  
 C:Keywords: neurotoxin

Query Match 16.8%; Score 411.5; DB 2; Length 1276;  
 Best Local Similarity 24.9%; Pred. No. 1.1e-17;  
 Matches 127; Conservative 96; Mismatches 164; Indels 123; Gaps 19;

OY 3 ESTPIPEFSYKNDGWDNEDIDVI-----LKKSTILNDINDNDIISDSCFNSSVIT 56  
 DB 836 FENTMPF-----NIFSYTNNSLKDIIIEYNFNSINDSKILSLONKKNALVDTSGYNAEVRV 891  
 OY 57 YPDQOLVGINCKATHL-----VNNSESEVIVHKAMDIENDMENFTVSPFLRVPK-VSA 111  
 DB 892 GDNVOL-----NTIYTNDFKLSGSGDKITVNNNLLYSAIYENSVSFWIKISKDTN 945  
 OY 112 SHLEQYGTNEYSTISSMKKHSLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLPDKF 171  
 DB 946 SH-----NEYTIINSIEGN-----SGMKLCIRNGNIEMTLQOVNRKYSKLIDYSESL 994  
 OY 172 NAYLANKWVEITITNDRSSANLYINGVLMSAEITGLAIREDNITLKLDRKNNNOY 231  
 DB 995 HTGYTNKWFVYITNNIMYKMLYINGELKQSKIDLELVKLDKTIYFGIDENIDENOM 1054  
 OY 232 VSIDKFRICFALKNEIKREIKLYTSYSTIFELDFWGNPLRYDTEYLIIPVASSKDVOLK 291  
 DB 1055 LMIKRFNIFSKELSDNEDIVIEGQILRVNIRKDYKWNPLKFDIEYIIN----- 1103  
 OY 292 NITDQWYLTNAPSYNGKINITY--RRLYNGLKFIKRYTPNNEIDSYKSGDFIKLYV 348  
 DB 1104 --DNYIDRYIAPE-SNVLVLYQPPDRSKLYTGNPITIKSVSDKNPY--SRLINGDNITLHM 1159  
 OY 349 STNNNEHIVGYRKDGNAFNNLDRILRVGNAPGIPLYKKMEAVKLRDLKT-YSVQ----- 402  
 DB 1160 LYSNRKYM-----LKKDKITIGCD--WTFVPTDEGW 460  
 OY 403 -----LKLYPD--KNASLGLVGTN-----GOIGNDENRDLIASN--WYFN-- 440  
 DB 1185 QNCYVALKQSNLNGNIGIGIFSIKNTVSKNKYCSQIFSSRRENTMLADYIKPWFSEKFN 1244  
 OY 441 -----HLKDKITIGCD--WTFVPTDEGW 460  
 DB 1245 AYTPAVATVYETKLLSTSSFEWKFISDPGW 1274

RESULT 14  
 S70582  
 botulinum neurotoxin type Dsa precursor - Clostridium botulinum phage d-sa  
 C:Species: Clostridium botulinum phage d-sa  
 A:Note: host Clostridium botulinum type D (strain South Africa)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Jun-2000  
 A:Accession: S70582  
 R:Morishita, K.; Koura, M.; Abe, N.; Fujii, N.; Fujinaga, Y.; Inoue, K.; Ogumad, K.  
 Biochim. Biophys. Acta 1307, 123-126, 1996  
 A:Title: Mosaic structures of neurotoxins produced from Clostridium botulinum types C  
 A:Reference number: S70582; MUID:96283801; PMID:8679691  
 A:Accession: S70582  
 A:Status: nucleic acid sequence not shown

A:Molecule type: DNA  
 A:Residues: 1-1285 <MOR>  
 A:Cross-references: EMBL:D38442; NID:q1374775; PIDN:BA07477.1; PID:q1374776  
 C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit new  
 C:disulfide bond. The heavy chain mediates the binding of toxin to the presynaptic membe  
 C:Superfamily: tetanus toxin  
 C:Keywords: disulfide bond; neurotoxin; transmembrane protein  
 F:1-447/Product: botulinum neurotoxin type Dsa light chain #status predicted <MAT1>  
 F:448-1285/Product: botulinum neurotoxin type Dsa heavy chain #status predicted <MAT2>

Query Match 16.5%; Score 406; DB 2; Length 1285;  
 Best Local Similarity 26.8%; Pred. No. 2.3e-17;

Matches 133; Conservative 87; Mismatches 190; Indels 86; Gaps 19;

```

QY 3 ESTPIPEFSYKMLDGVNEDIDVI-----LKKSTILNDINNDIISDGFNSVIT 56
Db 836 FENTIPF-----NIFSYTNLSLKDMINEFNSINDSKLSLQKKRTIAMDTSYNAEVRV 891
QY 57 YPDAGLV-----GNGKAIHLVNNESSEVIYHKAMIDYNDNFNFTVSFWLRVPR 108
Db 892 EGNVOLNPIPFDFELGSSG-----DDRGIIVYQENIYVAMAYESESISFWIRNK 944
QY 109 VSASHLEQYGTNEYSIISSMKHSISIGSGVSLKGNLITWTKDSAGEVRQITPR-DL 167
Db 945 -WVSNLP-----GYTIIIDSVKNN-----SGWSIGTISNPLVFTLKONENSEQDINFSDYI 993
QY 168 PDKFNAYLANKWFITTTNDRSLSANLYINGVLMGSAETIGLAIREDNNITLKDRCNN 227
Db 994 SKNAGY--NKFVFTYITNMGMNIYINGKLIDITIKVELTGINFSTIIFQMKIRN 1051
QY 228 NNQYVS-----IDKRIFCKALNPKKEIEKLYTSYSTITLRDQWGNPLKRYDEYLI 279
Db 1052 TGLTSDSDNINMMIRDFYIRAKELDDKDINLENSIQYTNVVKYDNGDLRYDEKYWI 1111
QY 280 PVASSSKDVQLKNTIDYMYLTNAPSYTNGKLNITYRRLNGELFKYIKRTYPPNEIDSEVK 339
Db 1112 NVNMYNRMYSKKG-----NGIYFNTRKNNNDPN---EGYKIIIRKIRGNTN-DTRVK 1159
QY 340 SGDFIKLYVSNNNNEHIYV-PRKQNAFNNIDRLIRVGYNAPGIDPYKKMEAVKRLDLKT 398
Db 1160 GENLVYNTITIDNKQYSLGMYKPSRNLGTDLVPLGALDQPMDEIRKYGSFIIQPCNTFDY 1219
QY 399 YSVQKLKYDDKNAS-IGD--VGTHNGOIGDNPNDIILASMWYFNH-----LKD----- 444
Db 1220 YASQLFLSSNATNTRLGLISIGSYSPFLGDD-----YWFNHEYLIPIYKTEHYAS 1269
QY 445 --KILGDMWFVPTDE 458
Db 1270 LLESTSTHWVFPASE 1285

```

RESULT 15  
 140817

botulinum toxin nontoxic component - Clostridium butyricum

C:Species: Clostridium butyricum

C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999

C:Accession: 140817

R:Fujii, N.; Yokosawa, N.; Oguma, K.; Yashiki, T.; Takeshi, K.; Isogai, E.;  
 Microbiol. Immunol. 37, 395-398, 1993

A:Title: Similarity in nucleotide sequence of the gene encoding nontoxic component of bc  
 Mashike.

A:Reference number: 140817; MUID:93360835; PMID:8355622

A:Accession: 140817

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-1162 <RES>

A:Cross-references: GB:012739; NID:9285709; PIDN:BA02231.1; PID:9285710

Query Match 11.0%; Score 269.5; DB 2; Length 1162;  
 Best Local Similarity 25.6%; Pred. No. 6.7e-09;

Matches 114; Conservative 67; Mismatches 144; Indels 121; Gaps 21;

QY 15 LDCWVDNEDIDVILKKSTILNDI----- 39

```

Db 749 LKQNTINTEKSHLLIMQNSFSLDFELDIQNMKNLFNLYTELLIEQNTSPYELSLAFQ 808
QY 40 --NNDIISDISEGNSSVITP-DAQLYPGINGKAIHLVNNESSEVIYHKAMIDYNDMFN 96
Db 809 EODNNVIDTSGKN-TLVEYFDIGLVGINNAHILTCANON--IKFTNDYFENGTLN 864
QY 97 NPTVSFWLRVPRVSASHLEQYGTNEYSIISSMKHSISIGS-----GWSYSLKGNLITW 151
Db 865 NESTYFWLR-----NLNQTIRSKL-----IGSKEDNGCWELYPENNGLVFN 906
QY 152 LKDSAGEVRQITPRDLPKFNAYLAN-----KWFYITITNDRSLSANL-YINGVLMGSAEI 206
Db 907 IIDSNGEK-----NYLSNISNKSWMHYIVISINRLKDQDLIIFIDNIVANEDI 955
QY 207 TGLGAIREDNNTITLKDRCNNNOQVYSIDKPRIFCKALNPKKEIEKLYTSYSTITLRDQW 266
Db 956 KEILNIVSSDIISLSD--NNVY--IEGLSVLAKTINSNELIDYFSDLNNSYIRNF 1010
QY 267 GNPRLRYDEY----YLIPVASSSKDVQLKNTIDYMYLTNAPSYTNGKLNITYRRLYNGLK 322
Db 1011 EETIQYNRKYTELFNYPPEIANKIEQNNNI---YLSN--NNENSINL-----FKPLK 1057
QY 323 FIKRYTPN-----NEIDSFVKSDFIKLYVSYNNE-HIVGPRKQGNAP--NN--L 369
Db 1058 FKLINTPNKQYQKMDVEVIFSVLDGTEKYLDISIDNRRIQLVNKNNAKYFTIINDIFI 1117
QY 370 DRILIRVGYNAPGIDPYKKMEAVKRLD 395
Db 1118 SNCLTLYNNVNVYL-----SIKMOD 1138

```

Search completed: December 1, 2002, 11:24:52  
 Job time : 54 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 10:23:47 ; Search time 25 Seconds

(without alignments)  
768.142 Million cell updates/sec

Title: US-09-816-467-2

Perfect score: 2456

Sequence: 1 MVFSTPIPFSTYSKNDQWVD.....DKILGDMYFVPTDEGTFND 463

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2451	99.8	1314	1	TETX_CLOTE
2	653.5	26.6	1295	1	BXAI_CLOBO
3	644.5	26.2	1296	1	BXA2_CLOBO
4	643.5	26.2	1295	1	BXA2_CLOBO
5	590	24.0	1290	1	BXE_CLOBO
6	565	23.0	1250	1	BXE_CLOBO
7	560.5	22.8	1250	1	BXE_CLOBO
8	517	21.1	1274	1	BXE_CLOBO
9	430	17.5	1290	1	BXC1_CLOBO
10	411.5	16.8	1276	1	BXD_CLOBO
11	269.5	11.0	1162	1	BXEN_CLOBO
12	263	10.7	1162	1	BXEN_CLOBO
13	244.5	10.0	1196	1	BXCN_CLOBO
14	134.5	5.5	1225	1	VG12_CVPR8
15	131	5.3	755	1	P100_HSV7J
16	128	5.2	987	1	YD94_METJA
17	128	5.2	1225	1	VG12_CVPR1
18	125	5.1	1449	1	VG12_CVPR1
19	123	5.0	1447	1	VG12_CVPR1
20	122.5	5.0	1447	1	VG12_CVPR1
21	119.5	4.9	2366	1	TOXB_CLODI
22	119	4.8	451	1	V110_FOMPV
23	119	4.8	1447	1	VG12_CVPR1
24	119	4.8	1447	1	VG12_CVPR1
25	118.5	4.8	1279	1	APU_THERY
26	117	4.8	1447	1	APU_THESA
27	117	4.8	1447	1	APU_THESA
28	116.5	4.7	634	1	YCX3_EUGGR
29	116	4.7	1138	1	C7AA_BACTU
30	116	4.7	1177	1	Y307_MYGE
31	115.5	4.7	839	1	YDDB_HABIN
32	115	4.7	993	1	NISB_LACLA
33	114	4.6	1729	1	RRP5_YEAST

34	113.5	4.6	613	1	CGAA_CLOBI
35	113.5	4.6	967	1	Y560_CLOAB
36	113.5	4.6	1127	1	MDM1_YEAST
37	113	4.6	1138	1	C7AB_BACUR
38	113	4.6	1449	1	VG12_CVPR5
39	112.5	4.6	1024	1	RPOB_PLAFA
40	112	4.6	806	1	SVPR_MYCCE
41	112	4.6	849	1	PHSG_SYNC3
42	112	4.6	1386	1	RPOD_MARPO
43	112	4.6	1956	1	ATX1_PLAFA
44	112	4.6	3144	1	VPI3_YEAST
45	111.5	4.5	592	1	Y036_HABIN

## ALIGNMENTS

RESULT 1  
TETX\_CLOTE  
ID TETX\_CLOTE STANDARD: PRT: 1314 AA.  
AC P04958:  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin).  
OS Clostridium tetani.  
OG Plasmid.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_Taxid=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87053814; PubMed=3536478;  
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,  
RA Weller U., Hudel M., Habermann E., Niemann H.;  
RT "Tetanus toxin: primary structure, expression in E. coli, and  
RT homology with botulinum toxins.";  
RL EMBO J. 5:2495-2502(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CM3911;  
RX MEDLINE=87040747; PubMed=3774547;  
RA Fairweather N.F., Lyness V.A.;  
RT "The complete nucleotide sequence of tetanus toxin.";  
RL Nucleic Acids Res. 14:7809-7812(1986).  
RN [3]  
RP SEQUENCE OF 742-1314 FROM N.A.  
RX MEDLINE=86085672; PubMed=3510187;  
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;  
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin  
RT fragment C in Escherichia coli.";  
RL J. Bacteriol. 165:21-27(1986).  
RN [4]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE=90201034; PubMed=2108021;  
RA Krieglstein K., Henschen A., Weller U., Habermann E.;  
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups  
RT in tetanus toxin.";  
RL Eur. J. Biochem. 168:39-45(1990).  
RN [5]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=92037649; PubMed=1935979;  
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;  
RT "Limited proteolysis of tetanus toxin. Relation to activity and  
RT identification of cleavage sites.";  
RL Eur. J. Biochem. 202:41-51(1991).  
RN [6]  
RP IDENTIFICATION AS ZINC-PROTEASE.  
RX MEDLINE=93010948; PubMed=1396558;  
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
RA Montecucco C.;  
RT "Tetanus toxin is a zinc protein and its inhibition of  
RT neurotransmitter release and protease activity depend on zinc.";

EMBO J. 11:3577-3583(1992).  
 [7]  
 IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=93063293; PubMed=1331807;  
 RA Scialvo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 Dasgupta B.R., Montecucco C.;  
 "Tetanus and botulinum B neurotoxins block neurotransmitter release  
 by proteolytic cleavage of synaptobrevin.";  
 Nature 359:832-835(1992).  
 RN  
 X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
 RX MEDLINE=97475217; PubMed=9334741;  
 RA Umhau T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
 Sax M.;  
 "Structure of the receptor binding fragment HC of tetanus  
 neurotoxin.";  
 Nat. Struct. Biol. 4:788-792(1997).  
 CC -I- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77  
 BOND OF SYNAPTOBREVIN-2.  
 CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN  
 SYNAPTOBREVIN.  
 CC -I- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
 YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
 AND ARE NON-TOXIC AFTER SEPARATION.  
 CC -I- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
 GANGLIOSIDE RECEPTORS.  
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC The European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as the content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-slb.ch/announce/  
 CC or send an email to license@isb-slb.ch).  
 CC -----  
 DR EMBL: X0436; CAA28033.1; -;  
 DR EMBL: M12739; AAA23282.1; -;  
 DR EMBL: X06214; CAA29564.1; -;  
 DR PIR: A25689; BTCLTN.  
 DR PDB: 1AP9; 29-APR-98.  
 DR PDB: 1ABD; 14-OCT-98.  
 DR MEROPS: M27.001; -;  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR000130; Zn\_Mpeptidase.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOTOXILYSIN.  
 DR PRODOM: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
 KW 3D-structure.  
 KM  
 FT INIT MET 0 0  
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.  
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.  
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 233 233 BY SIMILARITY.  
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT TRANSMEM 226 246 POTENTIAL.  
 FT TRANSMEM 669 689 POTENTIAL.  
 FT DISULFID 438 466 INTERCHAIN.  
 FT DISULFID 1076 1092  
 FT SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;  
 Query Match 99.8%; Score 2451; DB 1; Length 1314;  
 Best Local Similarity 100.0%; Pred. No. 4e-145;  
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 VNSTPIPFSTSKNDWCVDNEDIDVILKSTILNLDINNDIISDLSGFSNSVITTPDAQ 61

|||||  
 Db 653 VNSTPIPFSTSKNDWCVDNEDIDVILKSTILNLDINNDIISDLSGFSNSVITTPDAQ 912  
 QY 62 LVPGINGKAHLVNNSESEVYVHKAMDIEYNDMPNFVSWLPVPKYSASHLEQYGTNE 121  
 Db 913 LVPGINGKAHLVNNSESEVYVHKAMDIEYNDMPNFVSWLPVPKYSASHLEQYGTNE 972  
 QY 122 YSIISSMKKHSLSIGSGSVSLKGNLITWTLKDSAGEVROITFRDLPKFNAYLANKWVF 181  
 Db 973 YSIISSMKKHSLSIGSGSVSLKGNLITWTLKDSAGEVROITFRDLPKFNAYLANKWVF 1032  
 QY 182 ITTINDRLSSANLYINGVLMGSAEITGGAIREDDNITLTKDRCNNNOYVSTDKFRIFC 241  
 Db 1033 ITTINDRLSSANLYINGVLMGSAEITGGAIREDDNITLTKDRCNNNOYVSTDKFRIFC 1092  
 QY 242 KALNKEIEKLYTSLSTFTFLRDFWGNPLRDTFEEYLLIPVASSSDVOYLKNTDPMYLTN 301  
 Db 1093 KALNKEIEKLYTSLSTFTFLRDFWGNPLRDTFEEYLLIPVASSSDVOYLKNTDPMYLTN 1152  
 QY 302 APSYNGKLNITYRRLYGLKFIIRKYPNNEIDSFVKSQDFIKLYSVNNNEHIVGYPK 361  
 Db 1153 APSYNGKLNITYRRLYGLKFIIRKYPNNEIDSFVKSQDFIKLYSVNNNEHIVGYPK 1212  
 QY 362 DGNAPNNDRLIRVGYNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASIGLVGTNG 421  
 Db 1213 DGNAPNNDRLIRVGYNAPGIPLYKKMEAVKRLDKTYSVOLKLYDDKNASIGLVGTNG 1272  
 QY 422 QIGNDPNRDILIASNMVFNHLKDKILGCDWYFVPTDEGTND 463  
 Db 1273 QIGNDPNRDILIASNMVFNHLKDKILGCDWYFVPTDEGTND 1314  
 RESULT 2  
 BXAL\_CLOBO  
 ID BXAL\_CLOBO STANDARD: PRT: 1295 AA.  
 AC P10845; P18639; P01561;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)  
 DE Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-  
 chain; Botulinum neurotoxin A, heavy-chain].  
 GN BOTA OR BNA OR ATX.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 CC Clostridium.  
 OX NCBL\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCCT 2916;  
 RX MEDLINE=90235864; PubMed=2185020;  
 RA Thompson D.E., Brehm J.R., Outtram J.D., Swinfield T.-J.,  
 Ra Shone G.C., Atkinson T., Melling J., Minton N.P.;  
 "The complete amino acid sequence of the Clostridium botulinum type A  
 neurotoxin, deduced by nucleotide sequence analysis of the encoding  
 gene.";  
 Eur. J. Biochem. 189:73-81(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=62A;  
 RX MEDLINE=90264400; PubMed=2160960;  
 RA Binz B., Kiarzono H., Wille M., Frevent J., Wernars K., Niemann H.;  
 "The complete sequence of botulinum neurotoxin type A and comparison  
 with other clostridial neurotoxins.";  
 J. Biol. Chem. 265:9153-9158(1990).  
 RN [3]  
 RP SEQUENCE OF 1-65 FROM N.A.  
 RC STRAIN=62A;  
 RX MEDLINE=97016817; PubMed=8863443;  
 RA East A.R., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;  
 "Organization and phylogenetic interrelationships of genes encoding  
 components of the botulinum toxin complex in proteolytic Clostridium  
 botulinum types A, B, and F: evidence of chimeric sequences in the



RT gene encoding the nontoxic nonhemagglutinin component.";  
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).  
 RN [4]  
 RP SEQUENCE OF 1-34 FROM N.A.  
 RC STRAIN-Hall;  
 RX MEDLINE=89350959; PubMed=2669749;  
 RA Betley M.J., Somers E., Dasgupta B.R.;  
 RT "Characterization of botulinum type A neurotoxin gene: delineation of  
 the N-terminal encoding region.";  
 RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).  
 RN [5]  
 RP SEQUENCE OF 1-18 FROM N.A.  
 RC STRAIN-type A NIH;  
 RX MEDLINE=96096783; PubMed=8521962;  
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;  
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin  
 components of Clostridium botulinum type A progenitor toxins.";  
 RL FEBS Lett. 376:41-44(1995).  
 RN [6]  
 RP SEQUENCE OF 1-16.  
 RX MEDLINE=84178501; PubMed=6370252;  
 RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;  
 RT "Partial amino acid sequence of the heavy and light chains of  
 botulinum neurotoxin type A.";  
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).  
 RN [7]  
 RP SEQUENCE OF 1-46.  
 RX Dasgupta B.R., Foley J., Niece R.;  
 RT "Partial sequence of the light chain of botulinum neurotoxin type A.";  
 RL Biochemistry 26:4162-4162(1987).  
 RN [8]  
 RP SEQUENCE OF 1-5 AND 44-456.  
 RX MEDLINE=91120847; PubMed=2126206;  
 RA Dasgupta B.R., Dekleva M.L.;  
 RT "Botulinum neurotoxin type A: sequence of amino acids at the  
 N-terminus and around the nicking site.";  
 RL Biochimie 72:661-664(1990).  
 RN [9]  
 RP SEQUENCE OF 448-464 AND 872-895.  
 RX MEDLINE=89024662; PubMed=3178218;  
 RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;  
 RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two  
 halves and their partial sequences.";  
 RL Arch. Biochem. Biophys. 266:142-151(1988).  
 RN [10]  
 RP SEQUENCE OF 448-482.  
 RX MEDLINE=85285016; PubMed=3896784;  
 RA Shone C.C., Hambleton P., Melling J.;  
 RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin  
 and purification of two tryptic fragments. Proteolytic action near  
 the COOH-terminus of the heavy subunit destroys toxin-binding  
 activity.";  
 RL Eur. J. Biochem. 151:75-82(1985).  
 RN [11]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94063091; PubMed=824676;  
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,  
 RA Benfenati F., Wilson M.C., Montecucco C.;  
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct  
 COOH-terminal peptide bonds.";  
 RL FEBS Lett. 335:99-103(1993).  
 RN [12]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94124495; PubMed=8294407;  
 RA Blinz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,  
 RA Jahn R., Niemann H.;  
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";  
 RL J. Biol. Chem. 269:1617-1620(1994).  
 RN [13]  
 RP MUTAGENESIS OF GLU-261, PHE-265 AND TYR-365.  
 RX MEDLINE=21556941; PubMed=11700044;  
 RA Rigoni M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;  
 RT "Site-directed mutagenesis identifies active-site residues of the

RT Light chain of botulinum neurotoxin type a.";  
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).  
 RX MEDLINE=98455071; PubMed=9783750;  
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;  
 RT "Crystal structure of botulinum neurotoxin type A and implications  
 for toxicity.";  
 RL Nat. Struct. Biol. 5:898-902(1998).  
 CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin  
 binds with high affinity to peripheral neuronal presynaptic  
 membrane, is then internalized by receptor-mediated endocytosis.  
 CC The C-terminus of the heavy chain (H) is responsible for the  
 adherence of the toxin to the cell surface while the N-terminus  
 mediates transport of the light chain from the endocytic vesicle  
 to the cytosol. After translocation, the light chain (L)  
 hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking  
 neurotransmitter release. Inhibition of acetylcholine release  
 results in flaccid paralysis, with frequent heart or respiratory  
 failure.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 CC detected action on small molecule substrates.  
 CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
 CC heavy chain (H).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- PHARMACOLOGICAL: Available under the name BOTOX (Allergan) for  
 CC the treatment of strabismus and blepharospasm associated with  
 CC dystonia and cervical dystonia. Also used for the treatment of  
 CC hemifacial spasm and a number of other neurological disorders  
 CC characterized by abnormal muscle contraction.  
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of  
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -1- DATABASE: NAME=BOTOX product information Web site;  
 CC WWW="http://www.botox.com/index.jsp?hpproductinfo".  
 CC -1- DATABASE: NAME=Protein Spotlight;  
 CC NOTE=Issue 19 of February 2002;  
 CC WWW="http://www.expasy.org/spotlight/articles/spl1019.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52066; CAA36289.1; -  
 CC EMBL: M30196; AAA33262.1; -  
 CC EMBL: X92973; CAA63551.1; -  
 CC EMBL: D67030; BAA11051.1; -  
 CC EMBL: M27892; AAA23269.1; -  
 CC FIR: A35294; BTCLAB.  
 CC PIR: S09492; S09492.  
 CC PDB: 3BTA; 01-OCT-99.  
 CC MEROPS: M27.002; -  
 CC InterPro: IPR000395; Bontoxilysin.  
 CC InterPro: IPR000130; Zn\_MTPeptidse.  
 CC Pfam: PF01742; Peptidase\_M27; 1.  
 CC PRINTS: PR00760; BONTOTOXILYSIN.  
 CC ProDom: PD001963; Bontoxilysin; 1.  
 CC PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;  
 CC Pharmacological; 3D-structure.  
 CC INIT\_MDT 0  
 CC CHAIN 1 447  
 CC CHAIN 448 1295  
 CC METAL 222 223  
 CC METAL 223 223  
 CC METAL 226 226  
 CC METAL 261 261  
 CC DISULFID 429 453  
 CC BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.  
 CC BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.  
 CC ZINC (CATALYTIC).  
 CC ZINC (CATALYTIC).  
 CC ZINC (CATALYTIC).  
 CC INTERCHAIN.

```

FT DISULFID 1234 1279 POTENTIAL.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675
FT VARIANT 26 26 V -> A.
FT MUTAGEN 261 261 E->A: DRASTIC DECREASE IN ENZYMAIC
FT MUTAGEN 265 265 ACTIVITY
FT MUTAGEN 365 365 F->A: DECREASE IN ENZYMAIC ACTIVITY.
FT MUTAGEN 479 479 Y->A: DECREASE IN ENZYMAIC ACTIVITY.
FT CONFLICT 1 1 P -> Q (IN REF. 1).
FT CONFLICT 479 479 E -> P (IN REF. 9).
FT CONFLICT 875 875 T -> L (IN REF. 8).
FT CONFLICT 891 891 S -> K (IN REF. 8).
SQ SEQUENCE 1295 AA; 149322 MW; 858342F75486579 CRC64;

Query Match 26.6%; Score 653.5; DB 1; Length 1295;
Best Local Similarity 33.1%; Pred. No. 3.2e-33;
Matches 163; Conservative 88; Mismatches 159; Indels 83; Gaps 18;

OY 4 STPIPEYSKMLDGVNEDIDYILK-----KSTILINDINDIISDLSGFSNVITY 57
D 845 STDIPFOLSK---YVNOQLLSTFTEYIKNIITWTSTILNYESNHLIDLSRVASKIWIIG 900
OY 58 PDAOLVPGINGKAHLVNESSEVIVHKAMDIENFMNFTVSFWLRVPRK---VSASH 113
D 901 SKVNPDP-IDKNQIQFLBESSKTEVILKNAIVYNSMWFSTSEFWIRIPKFNISILN- 958
OY 114 LEQYGTNEYSTIISMKKHSISGWSVSLKGNLIWTLSKASAGEVQITFR-----DLP 168
D 959 -----NEYTIINCENN-----SGWKVSLNYGELIWTLODTQELIKRVVPRKYSOMINIS 1007
OY 169 DKFNAYILANKVAVITITDRTSALNYINGVLMSAETTGATREDNITLKLDRCNNN 228
D 1008 D-----YI-LNMEVITITNRLNNSKIYINGRLIDOKPLSNLGNHASNINMEFKLDGCDT 1062
OY 229 NOYVIDKFRICKALNPKIEKLTYSITFLDEFGNPLRYDEYVLIIPVASSKDV 288
D 1063 HRYIWKVFNLFDEKLEKEIKLDYDQNSGILKDFWGDYQYDKPYMLNLYPKNRYV 1122
OY 289 OLKN--ITDYMTLTPASPYTCKNLNIYR-RLXNGLKFLIKRYTPNNELDSFVKSDEIK 345
D 1123 DVNNNGIGGYML-KGPRGSVMTIYINSSLYRGTKFLIKRYASGKN-DNIVRNNDRYV 1180
OY 346 LVSYNNNEHLYGPRKDNAFNNDRLRVGYNAPGIPLYKKMEAVKLRLDK----- 397
D 1181 INVYVKNNEY-----RLATNASQAGYEKLTSLALEIPDVGNLSQVYVM 1222
OY 398 -----TYSYOLKLYDDKNASLGLVGTGNGQIGNDPNDRLILASNWTFNHLK--DKIL 447
D 1223 KSKNDGJITNCKKMLDQNDNGNDIGIFPH--QFNNIK--LVASNMVNRQIERSSRTL 1277
OY 448 GDDMYFVPTDEGM 460
D 1278 GCSWEFIPVDGM 1290

RESULT 3
BKG_CLOBO STANDARD; PRT: 1296 AA.
AC Q60393;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (Ec 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;

```

```

RA Campbell K., Collins M.D., East A.R.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentineense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins.";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAPs or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: secreted (by similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X74162; CAA52275.1; .
CC HSP: P10845; 3BTA.
CC HEROPS: M27.002; -.
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_mtpetidase.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONT0XILYSIN.
CC PRODOM: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC DR Neurotoxin; Hydrolase; Metalloprotease; Zinc.
CC INTI_MET 0
CC CHAIN 1 441
CC CHAIN 442 1296
CC METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT SITE 230 230 BY SIMILARITY.
CC METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; DC8E47E15F65C31 CRC64;

Query Match 26.2%; Score 644.5; DB 1; Length 1296;
Best Local Similarity 33.2%; Pred. No. 1.2e-32;
Matches 161; Conservative 82; Mismatches 183; Indels 59; Gaps 17;

OY 7 IFPS---YSKN-----LDCWVDNEDIDVILKSKTILNLDINDIISDLSGFSNVITY 57
D 840 IPFDLSLYTKRDTILIOVFNNYISN-----ISSNKLISLRGGLIDSSGIGATMNG 892
OY 58 PDAOLVPGINGKAHLVNESSEVIVHKAMDIENFMNFTVSFWLRVPRKVSASHLEQY 117
D 893 SDV-IFNDIGNQOFKLNNSNMTAHQSKFYVDSMFNFESINFWPRKYNNDIQY 951
OY 118 GNEYSIISMKKHSISGWSVSLKGNLIWTLSKASAGEVQITFR-DLPKFNAYLA 176
D 952 LQNEYTIISCIKN-----DSGKVSISIKGRRIITWLLIDVNAKSIFFESYIKNDISDYI- 1005
OY 177 NKWVFITINDRSLNLTNGVLMGSAETTGATREDNITLKLDRCNNNNOYVSIK 236
D 1006 NKWFSITINDRGNANITYINSLSKSEKILNDRNNSMDIDFKILNCIDTTRFWIKD 1065
OY 237 FRIFKALNPKIEKLTYSITFLDEFGNPLRYDEYVLIIPVASSKDVOLKNITDY 296
D 1066 FNFIFGRELNATEVSSLYWQSSNTTKDFWGNPLRYDQYLF--NQGQNIYIKVFSKA 1123

```



ID BXR.CLOBO STANDARD: PRT: 1290 AA.  
 AC P10844; P10843; (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)  
 DE (Bontolysin B).  
 GN BoNT.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 CX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92384550; PubMed=1514783;  
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,  
 RA Minton N.P.;  
 RA "Molecular cloning of the Clostridium botulinum structural gene  
 RA encoding the type B neurotoxin and determination of its entire  
 RA nucleotide sequence.";  
 RA Appl. Environ. Microbiol. 58:2345-2354(1992).  
 RL [2]  
 RN SEQUENCE OF 35-245 FROM N.A.  
 RP STRAIN=NCCTC 7273;  
 RC Szabo E.A., Pemberton J.M., Desmarchelier P.M.;  
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RN SEQUENCE OF 633-993 FROM N.A.  
 RP STRAIN=NCCTC 7273;  
 RC MEDLINE=94013372; PubMed=8408542;  
 RA Campbell K., East A.K., Collins M.D.;  
 RA "Gene probes for identification of the botulin neurotoxin gene and  
 RA specific identification of neurotoxin types B, E, and F.";  
 RA J. Clin. Microbiol. 31:2255-2262(1993).  
 RL [4]  
 RN SEQUENCE OF 1-44 AND 441-466.  
 RP STRAIN=657;  
 RC MEDLINE=89000987; PubMed=3139097;  
 RA Dasgupta B.R., Datta A.;  
 RA "Botulinum neurotoxin type B (strain 657): partial sequence and  
 RA similarity with tetanus toxin.";  
 RA Biochimie 70:811-817(1988).  
 RL [5]  
 RN SEQUENCE OF 1-16 AND 441-458.  
 RP STRAIN=OKRA;  
 RC MEDLINE=85197963; PubMed=3888113;  
 RA Schmidt J.J., Satyamourthy V., Dasgupta B.R.;  
 RA "Partial amino acid sequences of botulinum neurotoxins types B and  
 RA E.";  
 RA Arch. Biochem. Biophys. 238:544-548(1985).  
 RL [6]  
 RN IDENTIFICATION AS ZINC-PROTEASE.  
 RP MEDLINE=93054694; PubMed=1428690;  
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;  
 RA "Botulinum neurotoxins are zinc proteins.";  
 RA J. Biol. Chem. 267:23479-23483(1992).  
 RN [7]  
 RN IDENTIFICATION OF SUBSTRATE.  
 RP MEDLINE=93063293; PubMed=1331807;  
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
 RA Dasgupta B.R., Montecucco C.;  
 RA "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
 RA by proteolytic cleavage of synaptobrevin.";  
 RA Nature 359:832-835(1992).  
 RL -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF  
 SYNAPTOSOMAL-2.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO

CC detected action on small molecule substrates.  
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,  
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL  
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M81186; AAA2321.1; -;  
 DR EMBL: Z11934; CAA77991.1; -;  
 DR EMBL: X70817; CAA50148.1; -;  
 DR PIR: S07128; S07128.  
 DR PIR: S07155; S07155.  
 DR PIR: S08562; S08562.  
 DR PIR: S08573; S08573.  
 DR PIR: S08574; S08574.  
 DR PIR: A48940; A48940.  
 DR HSSP: P10845; 3BTA.  
 DR MEROPS: M27.002; -;  
 DR InterPro: IPR000395; Bontolysin.  
 DR InterPro: IPR000130; Zn\_Mpeptidse.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOLYSIN.  
 DR PRODOM: PD001963; Bontolysin.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 KW INIT\_MET 0  
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.  
 FT CHAIN 1 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.  
 FT METAL 229 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 230 230 BY SIMILARITY.  
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).  
 FT CONFLICT 29 29 T->M (IN REF. 4).  
 FT CONFLICT 217 217 R->G (IN REF. 2).  
 FT CONFLICT 224 224 A->S (IN REF. 2).  
 FT CONFLICT 463 463 S->R (IN REF. 4).  
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;  
 Query Match 24.0%; Score 590; DB 1; Length 1290;  
 Best local Similarity 31.7%; Pred. No. 2, 9e-29;  
 Matches 140; Conservative 89; Mismatches 184; Indels 28; Gaps 12;  
 QY 34 LINDLNNDLISDSEFNSVITYPPDAOLVPGINGK-AIHLVNNSESEVIVKAMDIEYN 92  
 DB 864 LNLKRXKNDLIDSLGKAYEYDDEL-----MDKNQFLTSANSKINVTQNTIFN 919  
 QY 93 DMENFTVSFWLVRPVKVSASHEOYSTNEYSISSMKHSLSGSGSVSLKGNLIWTL 152  
 DB 920 SVFLDSVSFWIRIPKXKNDGIONTHNEYTIINCKMNN-----SGWKISIRGRRIIWTL 974  
 QY 153 KSGAGVRQITPR-DLPDFENAVLANKWVFITTPNRLSSANLYINCVLMSARITGGA 211  
 DB 975 IDINCKTSVFEYENYREDISEYI-NKWFVYITNN-LNNAKIYINCKLESNDIKDIRE 1032  
 QY 212 IREDNNITLKDRCNNNNNOYSIDKFRICAKLNPKIEKLYSYSLITELDFWGNPLR 271  
 DB 1033 VLANGELIIRKLOSDIDRTQFIMWKYSIFTELSSQSIERYKIQSYSEVLKDFWGNPLM 1092  
 QY 272 YTEYELLIPVASSKVOALKNITDYMALTNAPSYTNKLNIIYRRLNGLKFLIKRTPN 331  
 DB 1093 YNKEYYMFNAGNKNISYIKLKDSVGEILTRSKYNQNSKYINRDLTIGKEFLIRKSN 1152

QY 332 NEI--DSFVKSDFEIKL--YVSNNNNHHVGYPKDGNFAENNLRILRVGNABRIPLYKKME 389  
 Db 1153 QSIWDDIVRKEDYIYLDFPNNLOEMRYVTT---KYFKKEEKFLFLADISDFEYFNIQ 1200  
 QY 390 AVKLRDKITSVQLKLYDKNNA--LGLVGH---NGOIGNDPNRDLIASNMYFNHLKD 444  
 Db 1209 IKEDDEPTSCQLLFKKDEESTDEIGIGHREYESCIVPEEYKDYCIKSWIYAEKVR 1266  
 QY 445 K---ILGCDWYFVPTDEGMT 461  
 Db 1269 KPYNLKKGCMWQFIKPKDEGMT 1289  
 RESULT 6  
 BXE\_CLOBO  
 ID BXE\_CLOBO STANDARD: PRT: 1250 AA.  
 AC 000496;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)  
 DE (Bontoxilysin E).  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Beluga;  
 RX MEDLINE=92181428; PubMed=1543481;  
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;  
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium  
 botulinum type E (strain Beluga) and Clostridium butyricum (strains  
 ATCC 43181 and ATCC 43755).";  
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92174922; PubMed=1541280;  
 RA Whelan S.M., Elmore M.J., Bodsoworth N.J., Atkinson T., Manton N.P.;  
 RT "The complete amino acid sequence of the Clostridium botulinum type-E  
 neurotoxin, derived by nucleotide-sequence analysis of the encoding  
 gene.";  
 RL Eur. J. Biochem. 204:657-667(1992).  
 RN [3]  
 RP SEQUENCE OF 1-251 FROM N.A.  
 RX MEDLINE=90264400; PubMed=2160960;  
 RA Blinz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;  
 RT "The complete sequence of botulinum neurotoxin type A and comparison  
 with other clostridial neurotoxins.";  
 RL J. Biol. Chem. 265:9155-9158(1990).  
 RN [4]  
 RP SEQUENCE OF 1-13.  
 RX MEDLINE=85197963; PubMed=388113;  
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;  
 RT "Partial amino acid sequences of botulinum neurotoxins types B and  
 E.";  
 RL Arch. Biochem. Biophys. 238:544-548(1995).  
 RN [5]  
 RP SEQUENCE OF 419-426.  
 RX MEDLINE=90344918; PubMed=2116911;  
 RA Gimenez J.A., Dasgupta B.R.;  
 RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C  
 reveals the site trypsin nicks and homology with tetanus  
 neurotoxin";  
 RL Biochimie 72:213-217(1990).  
 RN [6]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94063091; PubMed=8243676;  
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,  
 RT Benfenati F., Wilson M.C., Montecucco C.;  
 RL "Botulinum neurotoxins serotypes A and C cleave SNAP-25 at distinct  
 COOH-terminal peptide bonds.";  
 RL FEBS Lett. 335:99-103(1993).

```

RP [7] IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=9294407;
RA Jinn T., Blaszi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA Janz R., Niemann H.;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC 181 BOND IN SNAP-25.
CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62089; CAA43999.1; -.
DR EMBL; X62683; CAA44558.1; -.
DR PIR; A60027; A60027.
DR PIR; B35294; B35294.
DR PIR; JH0257; JH0257.
DR PIR; S08575; S08575.
DR PIR; S18111; S18111.
DR PIR; S21178; S21178.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mpeptidase.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin.1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc.
FT CHAIN 1 421
FT METAL 211 211
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT CONFLICT 176 176
FT CONFLICT 197 197
FT CONFLICT 339 339
FT CONFLICT 772 772
FT CONFLICT 962 963
FT CONFLICT 966 966
FT CONFLICT 1194 1194
SO SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;

Query Match 23.0%; Score 565; DB 1; Length 1250;
Best Local Similarity 29.7%; Pred. No. 9,9e-28;
Matches 146; Conservative 83; Mismatches 159; Indels 104; Gaps 17;

```

```

Db 822 IPFKLSYTD-----DKLLSYFNKFFRIKSSSVLNMRYKNDKYVDTSYDSSMIN 872
QY 56 TYPPQALVPGINGKAIHLVNNSSSEVYIHKAMDIEYNMFNFTVSEFWLRVKSASHLE 115
Db 873 INGVYKYP-TNRKNQFGIYNDLSEVNISQNDYIILDKNKNFESISFWRIPIYNTKIVN 931
QY 116 QYGTWEYSIISMKKHSLSIGSGSVSLKGNLITWLKDSAGEVQIOTFR-DLPKFNAY 174
Db 932 V--NNEYTIINCRDNN-----SGWKVSLNHNELIWTPEENRNGIKLAFVNGNANGISDY 985
QY 175 LANKWVFITTDRLSSANLYNGVLMGSAELTTCGATREDNNITLKDRCNNNNQYVSI 234
Db 966 I-NKMFVITTDRLSDSKLYINGNLIDKSLINLGNHVSQNLIFKLYNC-SYTRYIGI 1043
QY 235 DKFRICFALKNEKTEKLEKTYSTFLRDFGPNPLRYDTEYLLVPAVSSS-----KD- 287
Db 1044 KYENIFDKELDEIOTLTLSNEPNILKDFGNYLLVKEVYLLNVLKPNPFDKRKDS 1103
QY 288 -VOLKNITDYMYLTNAPSYNGKLNITYRRLYNGKLEIKRTYPPNNEIDSEYKSGDFITL 346
Db 1104 TISINNIRSTILLAN-----RLYSGIKVKIQRVNNSTNDNLVKKND--QV 1147
QY 347 YSY-----NNNEHYGYPKDGNFNNLDRILRVGNAPGIFLYKKMEA 390
Db 1148 YINFAVSKTHLPYADYATFTNKEKTIKISSGGRFNOV-----VYMSVG 1193
QY 391 VKLRDLKTYSVOLKLYDKNASLGLVGTNGIGDNDPNDLILASWYFNHLEKD--KILG 448
Db 1194 -----NCTNPFKNNGNNNIGLIGF-----KADTVVASTWYVYHMDRHTNSNG 1235
QY 449 CDMYFVPTDEGW 460
Db 1236 CFMNFISEHGW 1247

RESULT 7
BXE.CLOBU STANDARD; PRT: 1250 AA.
ID 1 BXE.CLOBU STANDARD; PRT: 1250 AA.
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontolysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1492;
OX 11
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 43181, and ATCC 43755;
RC MEDLINE=92181428; PubMed=1543481;
RX Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.:
"Sequences of the botulinum neurotoxin E derived from Clostridium
botulinum type E (strain Beluga) and Clostridium butyricum (strains
ATCC 43181 and ATCC 43755).";
RT Blochem. Biophys. Res. Commun. 183:107-113(1992).
RL [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
Yokosawa N., Yashiki T., Oguma K.:
"Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RA Gimenez J., Foley J., Dasgupta B.R.:
"Neurotoxin type E from Clostridium botulinum and C. butyricum:
partial sequence and comparison.";
RT FASEB J. 2:A1750-A1750(1988).

```

```

CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
CC EMBL: X62088; CAA43998.1; -
CC EMBL: X53180; CAA37321.1; -
CC PIR: JH0256; JH0256.
CC PIR: S16145; S16145.
CC HSSP: P10845; 3BTA.
CC MEROPS: M27.002; -.
CC InterPro: IPR000395; Bontolysin.
CC InterPro: IPR000130; Zn_MTPeptidse.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONTOLYSIN.
CC PRODOM: PD001963; Bontolysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC KW INIT_MET 0
CC FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
CC FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
CC FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 212 212 BY SIMILARITY.
CC FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT DISULFID 411 425 INTERCHAIN (PROBABLE).
CC FT CONFLICT 229 229 K -> M (in Ref. 2).
CC SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C312857 CRC64;

Query Match 22.8%; Score 560.5; DB 1; Length 1250;
Best Local Similarity 29.3%; Pred. No. 1.9e-27;
Matches 147; Conservative 83; Mismatches 164; Indels 107; Gaps 17;

QY 1 MVFST---PIPFYSKNDLCWVNEEDIDVL-----KSTLTNDINNDITSD 46
Db 813 NVIDLNNISIPFKLSYTD-----DKLLSYFNKFFRIKSSSVLNMRYKNDKYVD 863
QY 47 ISGNSSVITTPDAOLVPGINGKAIHLVNNSSSEVYIHKAMDIEYNMFNFTVSEFWLRV 106
Db 864 TSGYDSNININGDYKYP-TNRKNQFGIYNDLSEVNISQNDYIILDKNKNFESISFWRI 922
QY 107 PKVASHEQYGTNEYSIISMKKHSLSIGSGSVSLKGNLITWLKDSAGEVQIOTFR- 165
Db 923 PNYDKRYANV--NNEYTIINCRDNN-----SGWKVSLNHNELIWTPEENRNGIKLAFV 976
QY 166 DLPKFNAYLANKWVFITTDRLSSANLYNGVLMGSAELTTCGATREDNNITLKDRC 225
Db 977 GNANGISDYI-NKMFVITTDRLSDSKLYINGNLIDKSLINLGNHVSQNLIFKLYNC 1035
QY 226 NNNQYVSDIKFRICFALKNPFKEIKLYSYSTFLRDFGPNPLRYDTEYLLVPAVSS-- 283
Db 1036 -SYTRYIGIRYNIDFKELDEIOTLTLSNEPNANILKDFGNYLLVKEVYLLNVLKPN 1094

```

```

QY 284 -----SSKDVOLKNITDYMVLNAPSYTNGKLNLYRRLYNGLFKIRRYTPNNEIDSF 337
DB 1095 NFNRIRDSSTLSINNISTILLAN-----RLXSGIKVKIQRVNWSYTNL 1140
QY 338 VKSGDFELKLVSY-----NNNEHIVGPKRGNMFNNDRILIRVGNAPG 381
DB 1141 VRKND--QVYINFAVSKTHLLPLVADTATTNKKTKTIRSSGNNRNOV-----VYVNSVG 1193
QY 382 IPLYKKEAVKRLDKTYSVOLKLYDDKNASLGVTGHNQIGNDPNRDLIASWYFNH 441
DB 1194 -----NCTMNFKNNNNGNIGLLGF-----KADYVASTWYTYTH 1226
QY 442 LKDKI--LGCDWYFVPTDEGM 460
DB 1227 MRDNTNSNGFFWNFISEHGM 1247

RESULT 8
BXF_CLOBO
ID BXF_CLOBO STANDARD; PRT: 1274 AA.
AC P3096;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE (Bontoxilysin F).
GN BOPF.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RA MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RA "Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum.";
RT FEWS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=Hobbs FT10;
RA MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RA "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RA MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RA "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RA MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blas J., Link E., Cornille F.,
RA Roques B., Eykse E.M., Suedhof T.C., Jahn R., Niemann H.;
RA "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulin neurotoxins and tetanus toxin.";
RN [5]
RP F. Biol. Chem. 269:12764-12772(1994).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
CC BOND OF SYNAPTOBREVIN-1 AND -2..
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO

```

```

CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL; M92906; AAA23263.1; -
DR EMBL; S73676; AAC60475.1; -
DR EMBL; X70820; CAA50151.1; -
DR EMBL; X70816; CAA50147.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MRPepidse.
DR Pfam; PF01742; Peptidase_M27, 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FW CHAIN 1 436
FT CHAIN 1 437 1274
FT METAL 227 227
FT ACT_SITE 228 228
FT METAL 231 231
FT DISULFID 429 445
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 21.1%; Score 517; DB 1; Length 1274;
Best Local Similarity 29.1%; Pred. No. 9,9e-25;
Matches 131; Conservative 99; Mismatches 160; Indels 60; Gaps 17;

QY 29 LKSGTILNLDINNDIISGFSNVITYPPDAQVLGINKA-----HLVNNSS 79
DB 865 IKDSIILDMRENNKFIDISGNSI-----SINGNVIYISTNNRGITYNSRLS 914
QY 80 EYIYHKAMDIEYDNMFNFYSEWLKPVKSASLEOYGTN-EYSIISMKRHSLSIGSG 138
DB 915 EVNTAQNNDIYNSRYONFSISFWRIK---HYKPMNHREYTLINCGNANN---SG 966
QY 139 WSVSL---KGNLITWTLKDSAGEVQRIKPR-DLPKFNAYILANKVETITTDRLSSANL 194
DB 967 WKISLRVROCEIITWTDTSCKENLIFRIEELNRISYTI-NKMEIVITTNRLGNSRI 1025
QY 195 YINGVLGSAEITGILGAIREDNNITLKLDRGNMNNQYVSIIDKFRIFCKALNKEIEKYT 254
DB 1026 YINGNLIVKESISMGIHVDNLFLKIVGDDER-VYGIIRFKFENELDTTEITLIS 1084
QY 255 SYLSTIFLRQWGNPLKRDTEYIYLLPVASSKQVOKKITTDYMYTLNAPSYTNGKLNITY 314
DB 1085 NEPDPSILKNWGVLLYLNKRYLENLKRDKYITLN--SGILNINOORGVEGSEFLNY 1142
QY 315 RRLYNGLEKIFRYTP--NNEIDSEFKSGDPTIKLYVSNNEHIVGPKDGAFFNNDRI 372
DB 1143 -KLYEGEVALIRKGRPIDISWTDMFVKRND--AYI--NVYDRGEYRLYATKSEKEKI 1197
QY 373 LRYGYNAPGIPLYKKEAVKRLDKTYSVOLKLYDDKNASLGVTGHNQIGNDPNRDL 432
DB 1198 IRFS-----NLNDSIQIILYMDISGNCTNPNNGNSNIGLGFHSNN-----L 1242
QY 433 IASNWFENHLKDKIL--GCDWYFVPTDEGM 460

```

Db 1243 VASSWYNNIRNTSSNCEWSSISKENGW 1272

## RESULT 9

EXCL\_CLOBO STANDARD: PRT: 1290 AA.

AC P18640.  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-NOV-2002 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)  
 DE (Bontoxilysin C1).  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90370487; PubMed=2204031;  
 RA Hausner D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gill D.M.,  
 RT Boquet P., Popoff M.R.;  
 RL "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";  
 RL Nucleic Acids Res. 18:4924-4924(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Type C, Stockholm / C-ST;  
 RX MEDLINE=91024998; PubMed=2222445;  
 RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,  
 RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;  
 RT "The complete nucleotide sequence of the gene coding for botulinum  
 type C1 toxin in the C-ST phage genome.";  
 RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).  
 RN [3]  
 RP SEQUENCE OF 2-25.  
 RC STRAIN=Type C Stockholm / C-ST;  
 RX MEDLINE=88153072; PubMed=2450068;  
 RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,  
 RA Oguma K.;  
 RT "Establishment of a monoclonal antibody recognizing an antigenic site  
 common to Clostridium botulinum type B, C1, D, and E toxins and  
 tetanus toxin.";  
 RL Infect. Immun. 56:898-902(1988).  
 RN [4]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RX MEDLINE=94038966; PubMed=7901002;  
 RA Blas J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;  
 RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of  
 cleaving HPC-1/syntaxin.";  
 RL EMBO J. 12:4821-4828(1993).  
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 detected action on small molecule substrates.  
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
 HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,  
 WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL  
 FORMATION AND TOXIN BINDING, RESPECTIVELY.  
 CC -1- SUBCELLULAR LOCATION: secreted.  
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
 BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
 CC -1- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C  
 STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE  
 BACTERIOPHAGE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X66433; CAA47060.1;  
 DR EMBL: X72793; CAA51313.1;  
 DR EMBL: X53751; CAA37780.1;  
 DR EMBL: D90210; BAA14235.1;  
 DR EMBL: X62389; CAA44263.1;  
 DR PIR: S11291; S11291.  
 DR PIR: A35396; A35396.  
 DR PIR: A43503; A43503.  
 DR HSSP: P10845; 3BTA.  
 DR MEROPS: M27.002;  
 DR InterPro: IPR000395; Bontoxilysin.  
 DR InterPro: IPR000130; Zn\_Mpeptidase.  
 DR Pfam: PF01742; Peptidase\_M27; 1.  
 DR PRINTS: PR00760; BONTOXILYSIN.  
 DR ProDom: PD001963; Bontoxilysin; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 FT INIT\_MET 0  
 FT CHAIN 1 448  
 FT METAL 449 1290  
 FT METAL 228 228  
 FT ACT\_SITE 229 229  
 FT METAL 232 232  
 FT DISULFID 436 452  
 FT CONFLICT 84 84  
 SQ SEQUENCE 1290 AA; 148734 MW; 71FB379F97129E8 CRC64;

Query Match 17.5%; Score 430; DB 1; Length 1290;

Best Local Similarity 26.4%; Pred. No. 2.6e-19; Matches 135; Conservative 90; Mismatches 172; Indels 114; Gaps 21;

QY 3 FSTPIPEFSYKNDQVDEEDIVT-----LKKSTILNLDINNDISISGFNSSVIT 56  
 DB 839 FONTIF-----NIFYSTNNSLDKIDINEYFNINNDISKILSRKKTLDVPTSYNVESE 894  
 QY 57 YPDQOLVP-----GINKKAHLVNNESSEYVHKRAMDIENDMFNFTVSFWLRVPK 108  
 DB 895 EGVQOLNPIFPDFPKLGSSG-----DRKVIYQONENIYVNSMESFSISFWIRNK 947  
 QY 109 VSASHLEOYGTNEYSISSMKKHSLSIGSWSVSLKNNLIWTLKDSAGEVQITFR-DL 167  
 DB 948 -WVSNLP-----GYTIIDSVKNN-----SGMSIGIISNPLVFTLKQNEDESGSINFSDI 996  
 QY 168 PDKFNAYLANKWPFITITNRLSSANLYINGVLMGSAEITGLAIRDNNITLKLDR--- 224  
 DB 997 SNNAPGY--NKKWFEVYVITNNMGNMKLIYNGKLDITLKVLELGINFSKTIITFEINKIPD 1054  
 QY 225 -----CANNNOYVSIIDKFRIFCKALNKEIEKLYTSLSITFLDFEWNPLRYDTEYLI 279  
 DB 1055 TGLITSSDNNIMWIRFYIFAKELDDKIDNLFNSLOQYNNVAVDYGNDLRKKEYVM 1114  
 QY 280 PVASSKDVOLKNTDYMVLYLNAFSYNGKLNITYRRLYN-----GLKFIKRTKTPNNEI 334  
 DB 1115 NI-----DYL-----NRYMANSROIYFNTRNNNDNEGKIIIKIRLGWTN- 1157  
 QY 335 DSPYKSGDFIKLYSYNNNEHIVCPDQGNANFNLDLRLRGVYAPGIPYKKKEAVK-- 392  
 DB 1158 DTRVRGDIILYFDMTINKAV-----NLEMKNETMYADNHSTEDIIYAIGLRQKIDI 1209  
 QY 393 -----LRDLKITYSVQ--KLYDDKNAS-LGLVGTNGQIGNDPNRDILASWY- 438  
 DB 1210 NDNITFQIQPNNNYYIASQIFKFNNGENISGICSTGYFRIGCD-----WYR 1259  
 QY 439 FNHLKDKI-----LGCDFYFVPTDE 458  
 DB 1260 HNYLPTVYKQGNVASLLESTSTHWGFVPSV 1290



RESULT 10  
 BND\_CLOBO STANDARD; PRT; 1276 AA.  
 ID BND\_CLOBO  
 AC P19321;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last annotation update)  
 DTR 15-JUN-2002 (Rel. 41, Last annotation update)  
 DDE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BONT/D)  
 DE (Bontolysin D).  
 GN BONT.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OC NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BVD/-3.  
 RA MEDLINE=91016853; PubMed=2216736;  
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,  
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;  
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum  
 RT neurotoxin type D.";  
 RL Nucleic Acids Res. 18:5556-5556(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CB16;  
 RA MEDLINE=93042276; PubMed=1420572;  
 RA Sunagawa H., Ohnaya T., Watanabe T., Inoue K.;  
 RT "The complete amino acid sequence of the Clostridium botulinum type D  
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding  
 RT phage d-16 phi genome.";  
 RL J. Vet. Med. Sci. 54:905-913(1992).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC STRAIN-D-SA, and D-1873;  
 RA MEDLINE=89339741; PubMed=2668193;  
 RA Morishita K., Syuto B., Kubo S., Oguma K.;  
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D  
 RT strains.";  
 RL Infect. Immun. 57:2886-2891(1989).  
 RN [4]  
 RP IDENTIFICATION OF SUBSTRATE.  
 RA MEDLINE=94230352; PubMed=8175689;  
 RA Yamasaki S., Baumeister A., Binz T., Blas J., Link E., Cornille F.,  
 RA Roques B., Eykne E.M., Suedhof T.C., Jahn R., Niemann H.;  
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and  
 RT F botulinum neurotoxins and tetanus toxin.";  
 RL J. Biol. Chem. 269:12764-12772(1994).  
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-1-LEU-61 BOND OF  
 CC SYNAPTOSOMAL-1 AND -2.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO  
 CC detected action on small molecule substrates.  
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,  
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIANE CHANNEL  
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF  
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
 CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D  
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE  
 CC BACTERIOPHAGE.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X54254; CAA38175.1; -  
 CC EMBL: S49407; AAB24244.1; -  
 CC PIR: S11455; S11455.  
 CC HSSP: P10845; 3BTA.  
 CC MEROPS: M27.002; -.  
 CC InterPro: IPR000395; Bontolysin.  
 CC InterPro: IPR000130; Zn\_Mpeptidase.  
 CC Pfam: PF01742; Peptidase\_M27; 1.  
 CC PRINTS: PR00760; BONTOLYSIN.  
 CC PRODOM: PD001963; Bontolysin.  
 CC PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 CC KEGG: Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 CC CHAIN 1 442  
 CC METAL 443 1276  
 CC ACT\_SITE 229 229  
 CC ACT\_SITE 230 230  
 CC METAL 233 233  
 CC DISULFID 437 450  
 CC VARIANT 15 16  
 CC VARIANT 17 18  
 CC VARIANT 452 452  
 CC VARIANT 457 457  
 CC VARIANT 457 457  
 CC VARIANT 462 462  
 CC VARIANT 489 489  
 CC VARIANT 644 644  
 CC VARIANT 1122 1122  
 CC SEQUENCE 1276 AA; 146871 MW; C1EC50F46C8233E2 CRC64;  
 Query Match 16.88; Score 411.5; DB 1; Length 1276;  
 Best Local Similarity 24.9%; Pred. No. 3.7e-18;  
 Matches 127; Conservative 96; Mismatches 164; Indels 123; Gaps 19;  
 QY 3 FSPPIPSYSKSLDNCWDNEDIDVI-----LKKSTLINDINDIISDGSFNSSVTT 56  
 DB 836 FENIMPF-----NIFSYNNLSLKDITINEYNSINDSKLSLQNKALVDVSGYNAEYV 891  
 QY 57 YPDAPGVINGKAIRH-----VNNSESEVIAHKAMDIYENMFNFVSPFMRVPRK-USA 111  
 DB 892 GDNVQL-----NTIYNDFLSSGGKLIYNNLNLILAYENSVSFWTKISKDLTN 945  
 QY 112 SHLEQYCTNEYSITSSKKKSHLSIGSGWSVSLKGNLITWLKDSAGEVQITFRDLPDKF 171  
 DB 946 SH-----NEYTIINSLEON-----SQMKLCIRNGINLEMIIDVVRKRYKSLFDYSESLS 994  
 QY 172 NAYLANKWVEITITNDRLSSANLYINGVLMGSAETTGAIREDNNITLKLDRCNNNOY 231  
 DB 995 HGTGTAKWFEVITITNMGIMKLYINDELKOSQKIEDLEVKLDTYFGIDENIDENOM 1054  
 QY 232 VSDKRRIFPKALNPKIEKLYTSYSITFLRDPWGNPLRYDTEYLLIPVASSSKDYOAK 291  
 DB 1055 LWRIDNIESKESLNEDINIVYEQILRWYIKDYGWPLKDETYIIN----- 1103  
 QY 292 NLUDYVLTNPAPSTNGKLIY---RLNLGKLTITKRIYPPNEIDSFVSGDFIKLYV 348  
 DB 1104 --DNYIDRYIAP-SNVLVVOYPPDRSKLYTGNPTIKSVSDKNPY-SRLINGNIIILHM 1159  
 QY 349 SYNNNEHIVGYPKDGNAFNMLDLRLRGVNAFGIPLTKKMEAYVLRLDKT-YSVO----- 402  
 DB 1160 LNSRKRYMI-----IROTDTIYAIOGGEGCS 1184  
 QY 403 -----LKLVD-KNASLGLVGTN-----GQIGDNDNRDILLASN---WYFN--- 440  
 DB 1185 QNCYVALKLOSNNGNGIGIFSIKNIYSSKNKYSQIFSSPRENTMLADYIKPWFSPFKN 1244  
 QY 441 -----HLKDKITGCD--WYFPTDEGW 460  
 DB 1245 ATPPAVAVNYETKLLSTSSFWKTSRDPGW 1274

```

RESULT 11
BXEN_CLOBO STANDARD; PRT: 1162 AA.
AC 006366;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Botulinum neurotoxin type E, nontoxic component.
OS Clostridium butyricum; Clostridia; Clostridiales; Clostridiaceae;
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Bl6340;
RX MEDLINE=9336083; PubMed=8355622;
RA Fujii N., Kimura K., Yokosawa N., Oguma K., Yashiki T.,
RA Takeshi K., Ohyama T., Isogai E., Isogai H.;
RT "Similarity in nucleotide sequence of the gene encoding nontoxic
RT component of botulinum toxin produced by toxigenic Clostridium
RT butyricum strain Bl6340 and Clostridium botulinum type E strain
RT Mashike.";
RT Microbiol. Immunol. 37:395-398(1993).
CC -I- FUNCTION: THE NONTXIC COMPONENT IS NECESSARY TO MAINTAIN
CC TOXICITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D12739; BAA02231.1;
CC InterPro: IPR000395; Bontoxilysin.
CC Pfam: PF01742; Peptidase_M27.1.
CC PRINTS: PR00760; BONTOTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin.1.
CC Neurotoxin.
CC SEQUENCE 1162 AA; 136829 MW; C86E9BE932DA78E4 CRC64;
Query Match 11.0%; Score 269.5; DB 1; Length 1162;
Best Local Similarity 25.6%; Pred. No. 2.3e-09;
Matches 114; Conservative 67; Mismatches 144; Indels 121; Gaps 21;
QY 15 LDCWVNEEDIDVILKSTILNDI----- 39
DB 749 LKCTNINTEKSHLIMQNSFNLDPFELDIQNMKNLFNLYTELLKEQTSPEYELSLAFQ 808
QY 40 --NNDIISDISGNSSVITYP-DAQLVPGINGKAHLVNNSESEVYVHKAMDIEYDMEN 96
DB 809 EODNNVYIGTSGKN-FLVEYPRDIGLVGINNAHLTGANON---IKFTNDFENGLTN 864
QY 97 NPTVSFRLVPKVSASHLBOYGTNEYISISMKHSLSIGS-----GWSVSLKGNLIWT 151
DB 865 NPSIYFWLR-----NLQNTIKSL-----IGSKEDNGWELFEFNGGLVEN 906
QY 152 LKDSAGEVROITFRDLPKFNAYLAN---KWVFTITNDRLSANL-YINGVLMSAEI 206
DB 907 IIDSNGNEK-----NLYLSINMSKSWHYIVISIRKEDOLLIFIDNIIIVANEDI 955
QY 207 TGGCAIREDDNITLKDRCNNNOVYSIDKFRIFCKALNPKIEKLYTSYSLTFPRDFW 266
DB 956 KELINTYSSDIISLSD--NNNVY--TEGLSVLNKTTINSNELITDYFSDLNNSYIRNFD 1010
QY 267 GNPRLRYDEV-----YLIPVASSKDVOLKNTIDYMYLTNAPSYTNGKLNITYRRLYNGLK 322
DB 1011 EELLQYNRTELELVYPELAIKIBQNNNI-----YLSN---NNNSLN-----FKPLK 1057
QY 323 FIKRRTYPN-----NEIDSVKSGDFIKLIYSYNNE-HIYGVPKDGNAF--NN---L 369

```

```

DB 1058 FKLLINTPNKOYVOKMBEIVSFVLDGTETKYLIDISDNNRIQDLYONKNNAKFIINNDIFI 1117
QY 370 DRILRGYNAPGIDPLYYKMEAVKLKD 395
DB 1118 SNCITLITNNVNVYL-----SIKNOD 1138
RESULT 12
BXEN_CLOBO STANDARD; PRT: 1162 AA.
AC P46082;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Botulinum neurotoxin type E, nontoxic component.
GN ENT-120.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Mashike;
RX MEDLINE=93195515; PubMed=8450310;
RA Fujii N., Kimura K., Yokosawa N., Yashiki T., Tsuzuki K., Oguma K.;
RT "The complete nucleotide sequence of the gene encoding the nontoxic
RT component of Clostridium botulinum type E progenitor toxin.";
RT J. Gen. Microbiol. 139:79-86(1993).
CC -I- FUNCTION: THE NONTXIC COMPONENT IS NECESSARY TO MAINTAIN
CC TOXICITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D12697; BAA02194.1;
CC InterPro: IPR000395; Bontoxilysin.
CC Pfam: PF01742; Peptidase_M27.1.
CC PRINTS: PR00760; BONTOTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin.1.
CC Neurotoxin.
CC SEQUENCE 1162 AA; 136856 MW; 96468EDDADF0F39D CRC64;
Query Match 10.7%; Score 263; DB 1; Length 1162;
Best Local Similarity 23.7%; Pred. No. 5.7e-09;
Matches 111; Conservative 76; Mismatches 136; Indels 146; Gaps 21;
QY 15 LDCWVNEEDIDVILKSTILNDI----- 39
DB 749 LKCTNINTEKSHLIMQNSFNLDPFELDIQNMKNLFNLYTELLKEQTSPEYELSLAFQ 808
QY 40 --NNDIISDISGNSSVITYP-DAQLVPGINGKAHLVNNSESEVYVHKAMDIEYDMEN 96
DB 809 EODNNVYIGTSGKN-FLVEYPRDIGLVGINNAHLTGANON---IKFTNDFENGLTN 864
QY 97 NPTVSFRLVPKVSASHLBOYGTNEYISISMKHSLSIGS-----GWSVSLKGNLIWT 156
DB 865 NPSIYFWLR-----NLQNTIKSL-----IGSKEDNGWELFEFNGGLVEN 911
QY 157 GEVROITFRDLPKFNAYLAN---KWVFTITNDRLSANL-YINGVLMSAEITSLGA 211
DB 912 GNKK-----NLYLSINMSKSWHYIVISIRKEDOLLIFIDNIIIVANEDIKELN 960
QY 212 IREDNNTTLKDCNNNOVYSIDKFRIFCKALNPKIEKLYTSYSLTFPRDFWGPLR 271
DB 961 IYSDIISLSD--NNNVY--TEGLSVLNKTTINSNELITDYFSDLNNSYIRNDEEILQ 1015
QY 272 YDREY-----YLIPVASSKDVOLKNTIDYMYLTNAPSYTNGKLNITYRRLYNGLKFIKR 327

```



FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 850 850 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 976 976 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1100 1100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1117 1117 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1134 1134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1147 1147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1225 AA: 134728 MW: 1D7650D89A48C9D CRC64;

Query Match 5.5%; Score 134.5; DB 1; Length 1225;  
 Best Local Similarity 18.1%; Pred. No. 0.61;  
 Matches 101; Conservative 85; Mismatches 180; Indels 193; Gaps 27;

QY 6 PTP-PSYKNDLCWDNED-----IDVIL-----KSTJLNDINND 42  
 DB 83 PPSFEAASTCEGADPDGCGAVLNTVDYIRENLFTTNVSGKATVSLTTGG 142  
 QY 43 IISDIGFN-----SSVITYPPDAQVLP-GI-NGKAIHLVNNSESEVIVHKADIEYNDME 95  
 DB 143 VTEICICYNDIYSDSSFSSTGE---IPFGVTNPRCYV-----LY 180  
 QY 96 NNFVTSFMLRVP---KVSASHLEOYGTMEYSITSSMKKHSISGWSVSLKGNLIMT 151  
 DB 181 NGALKYGLTPSVYKEIAISKWHEFYINGYNFSTFPIDCIS---FNLITGDSOVFWT 236  
 QY 152 LKASAEVQRITRDLRDKFNAYLANKWFEITTNDRLSANLYINGVLSAEITGLA 211  
 DB 237 I-----ATTSTYEALOYENTAITNVT-YCHSYV----- 264  
 QY 212 IREDNNITTLKDRCONNOVYSIDKFRIFCKALNPKIEKLYSYSLTFPL----- 262  
 DB 265 ---NNIKCSQULANLNGFYPSVSSSEV--GSYKNSYV--LLPSFLHTITVNTTIGLGMK 316  
 QY 263 RDMWGNPLRVDTEYLIPVASSSKDYOL-----KNITDYM- 297  
 DB 317 RSGGPIASTSLNITLPMQDNNITDYCVRSDFSVYVHSTCSALMDNFKRNCIDVLD 376  
 QY 298 ---YLTNASTYNGKLNITYRRLYN-----GLKEIIRKRYTPNNEIDSEVKS GD 342  
 DB 377 ATNAVIKGTCPFSFDKLNLYL--TFNFKCLSLSPVGANCKEFDVARTRTN--DOFVRS-- 430  
 QY 343 FIKLYSVYNNNEHIVGPKGNAFNNIDRLIRVG-----YANPGIPLYKKMAVAVLRD 395  
 DB 431 ---LYVIEEGDSIVGYPSONSGIHLDS--VLHDSCTDYNIGRTGVGIIIRQTNFTLLSG 486  
 QY 396 LKTYSVQLKLYDOKNASTGLV-----GTHNGOIGNPNEDILLASNM-- 437  
 DB 487 LYTYSLSGDLGFKNVSDDGYIYVTPCDVNSQAAILDGAIVGAL-TSINSLLALHTMTI 545  
 QY 438 -----YFNHLKDKILG 448  
 DB 546 TPNFYYSIVYNTDKTRG 564

RESULT 15  
 P100\_HSV7J STANDARD; PRT; 755 AA.  
 ID P100\_HSV7J  
 AC P52519;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Large structural phosphoprotein homolog (Pp100).  
 GN U11.

OS Human herpesvirus (type 7 / strain J1) (HHV7).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=57278;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nicholas J.;  
 RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO THE LARGE STRUCTURAL PHOSPHOPROTEINS OF HSV-6 AND  
 CC HCMV UL32.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: 043400; AAC54672.1;  
 KW Matrix protein: Phosphorylation.  
 SQ SEQUENCE 755 AA: 86580 MW: 4083744CC3F91DA CRC64;

Query Match 5.3%; Score 131; DB 1; Length 755;  
 Best Local Similarity 22.3%; Pred. No. 0.54;  
 Matches 110; Conservative 71; Mismatches 194; Indels 118; Gaps 27;

QY 44 ISD-----ISGFSSVITP--DAQVPGI-----NCKAIH--LVNNESEVIVHK 85  
 DB 11 ISDEAKCFLSREFENISLPEVDIRENPWLSOCIAKGTGINSINNVKTLNNLLMIYFHQ 70  
 QY 86 AM---DIEYNDMPNFTVSEFWLVPVVSASHLBO-YGTMEYSITSSMKKHSISGWSV 141  
 DB 71 TLCKKRPDVEWQEL-----LKVQKILKDYLRQRMITDYSLSFSNK-----VGFEF 119  
 QY 142 SLK-----GNLITWLKDSAGEYRQTFEDLP-----KENAY-LANK 178  
 DB 120 EFKNVAKDLKLKLSFLRMGTVAHADVNLTTERRAIGENLOKANNMLSTYIQLVDP 179  
 QY 179 W-----VEITTNDRLSANLYINGVLMGS---AEITGLAIREDDNITTLKDRCONNOY 231  
 DB 180 WENGYIVYVINKLILYGNLLI--TLHGSMMNEKIALNTINEKKNALIK--ATENKNE 235  
 QY 232 VSIDKFRIFCKALNPKIEKLYSYSLTFPLRDFWGNPLKVDTEYLIPVASSSKD----- 287  
 DB 236 VSIYSYQILSLPLTSHRY---TSFEKI-LTEDEP--DVITKSLHLALPVKSTWDDVKF 288  
 QY 288 ---VQ-LKNITDYMVLTNASTYNGKLNITYRRLYNGLKFEIIRKRYTPNNEI--DSFVKS 340  
 DB 289 TPBEIOEFKVLTDLSKSLSNQFSSSKTSHGSSFPEPIKTEQORSNNTLSKDLVGS 348  
 QY 341 GDFIKLYSVYNNNEHIVGPKDGNFNNIDRLIRVGYNAPGIPLYKKMAVAVLRDLYTS 400  
 DB 349 ED--GLSSVYKDSMILDEPRNSTSINSSKMHRI-----LQTEIIDLDTQTMHR 396  
 QY 401 VOLKLY-----DKNA---SLGVGHNGOIGNDPNEDILLASNM--KDK 445  
 DB 397 PEDKYNQFNEIAVAPDSINOVIDFLSKLDHNS-----NKVIDIVSPKVVAVQLKRNK 450  
 QY 446 ILGCDWYFVPTDE 458  
 DB 451 IDYHSTFPLENE 463

Search completed: December 1, 2002, 11:22:06  
 Job time : 35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2002, 11:16:52 ; Search time 80 Seconds

(without alignments)  
1192.498 Million cell updates/sec

Title: US-09-816-467-2  
Perfect score: 2456  
Sequence: 1 MVESTRIPFSYSKNLDCWVD.....DKILCCDWYVPDTEGWTND 463

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP\_ARCHAEA:\*
- 2: SP\_BACTERIA:\*
- 3: SP\_FUNGI:\*
- 4: SP\_HUMAN:\*
- 5: SP\_INVERTEBRATE:\*
- 6: SP\_MAMMAL:\*
- 7: SP\_MHC:\*
- 8: SP\_ORGANELLE:\*
- 9: SP\_PHAGE:\*
- 10: SP\_PLANT:\*
- 11: SP\_RODENT:\*
- 12: SP\_VIRUS:\*
- 13: SP\_VERTEBRATE:\*
- 14: SP\_UNCLASSIFIED:\*
- 15: SP\_VIRUS:\*
- 16: SP\_BACTERIAP:\*
- 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2359	96.1	1310	2	093N27
2	2340	95.3	451	2	091A13
3	617	25.1	1291	2	008077
4	600	24.4	1291	2	0933K0
5	599	24.4	1291	2	092A58
6	597	24.3	1291	2	093G71
7	571	23.2	1268	2	045851
8	570.5	23.2	1251	2	09X395
9	567	23.1	441	2	09X708
10	565	23.0	1255	2	09FAR6
11	553	22.5	1278	2	057236
12	532.5	21.7	1280	2	092A55
13	430	17.5	1291	2	093HT3
14	417.5	17.0	1280	2	091B57
15	417.5	17.0	1280	2	045849
16	411.5	16.8	1275	12	09QTC7

17	407	16.6	1285	2	091BR1	091br1 clostridium
18	406	16.5	1285	2	045967	045967 clostridium
19	274.5	11.2	1197	2	033871	033871 clostridium
20	274.5	11.2	1197	2	P71117	P71117 clostridium
21	269	11.0	1197	2	092A59	092a59 clostridium
22	262	10.7	1197	2	045888	045888 clostridium
23	253.5	10.3	1196	2	091BR2	091br2 clostridium
24	252.5	10.3	1196	2	091BR2	091br2 clostridium
25	246.5	10.0	1196	2	053550	053550 clostridium
26	246.5	10.0	1196	2	092X77	092x77 clostridium
27	246.5	10.0	1196	2	045893	045893 clostridium
28	245.5	10.0	1196	2	069276	069276 clostridium
29	244.5	10.0	1196	2	045914	045914 clostridium
30	244.5	10.0	1196	2	093HT4	093ht4 clostridium
31	244.5	10.0	1196	2	038197	038197 clostridium
32	243	9.9	1162	2	092A56	092a56 clostridium
33	241	9.8	361	2	045846	045846 clostridium
34	241	9.8	361	2	045848	045848 clostridium
35	239	9.7	1198	2	006018	006018 clostridium
36	228	9.3	1160	2	087710	087710 clostridium
37	225	9.2	1193	2	045914	045914 clostridium
38	225	9.2	1193	2	P71107	P71107 clostridium
39	224	9.1	1161	2	069276	069276 clostridium
40	222	9.0	1193	2	045880	045880 clostridium
41	220	9.0	1163	2	045850	045850 clostridium
42	217.5	8.9	367	2	045861	045861 clostridium
43	217.5	8.9	367	2	045862	045862 clostridium
44	217	8.8	1161	2	045891	045891 clostridium
45	217	8.8	1193	2	P71108	P71108 clostridium

## ALIGNMENTS

### RESULT 1

093N27 PRELIMINARY; PRT; 1310 AA.  
AC 093N27;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Tetanus toxin (Fragment).  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_Taxid=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shumin Z., Dianliang L.;  
RT "Cloning and sequence analysis of tetanus toxin gene."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF389424; AAK72964.2; -  
DR InterPro: IPR000395; Bontotoxin.  
DR InterPro: IPR001064; Crystallin.  
DR InterPro: IPR000130; Zn\_M7peptidase.  
DR Pfam: PF01742; Peptidase\_M27; 1.  
DR ProDom: PD001963; Bontotoxin.  
DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDG91418E450 CRC64;

Query Match 96.1%; Score 2359; DB 2; Length 1310;  
Best Local Similarity 98.0%; Pred. No. 8.5e-132;  
Matches 447; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 VESTRIPFSYSKNLDCWVDNEEDIDVILKSTLLNDINNDISDGSFSSVITTPDAQ 61  
DB 855 VESTRIPFSYSKNLDCWVDNEEDIDVILKSTLLNDINNDISDGSFSSVITTPDAQ 914  
QY 62 LVPGNGKAHILVNNESSFVTHKAMDIEYNMENNFTYSEFWLRVRKVASHLDEYGTNE 121  
|||||

```

Db 915 LVPGINGKAIHLVNESSEVIVHKAMDEYNDEMNFVFWLVRPVASSHLEQVPTNE 974
QY 122 YSIISSMKKSHLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLDPKFNAYLANKWV 181
   |||||
Db 975 YSIISSMKKSHLSIGSGWSVSLKGNLIMTLKDSAGEVROITFRDLDPKFNAYLANKWV 1034
QY 182 ITTINDRLSSANLIVNGVMSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFC 241
   |||||
Db 1035 ITTINDRLSSANLIVNGVMSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFC 1094
QY 242 KALNPKETEKLYTSLSTLFRDVGWGNPLRDYETLYLIPVASSKDVOLKNTTYMYLTN 301
   |||||
Db 1095 KALNPKETEKLYTSLSTLFRDVGWGNPLRDYETLYLIPVASSKDVOLKNTTYMYLTN 1154
QY 302 APSTYNGKLNITRYRLYGLKFEIIRKTPPNEIDSFVKSQDFIKLYSVNNNEHIVGYPK 361
   |||||
Db 1155 APSTYNGKLNITRYRLYGLKFEIIRKTPPNEIDSFVKSQDFIKLYSVNNNEHIVGYPK 1214
QY 362 DGNMFNNIDRLRVGNVAPGIPLYKKMEAVKLRDLKITYSVOLKLYDDKNASIGLVGTING 421
   |||||
Db 1215 DGNMFNNIDRLRVGNVAPGIPLYKKMEAVKLRDLKITYSVOLKLYDDKNASIGLVGTING 1274
QY 422 QIGNDPNRDIILASNMVYNNHLKDKITLGCDDWYFVPTD 457
   |||||
Db 1275 QIGNDPNRDIILASNMVYNNHLKDKITLGCDDWYFVPTD 1310

```

## RESULT 2

```

09LAI3 PRELIMINARY: PRT: 451 AA.
ID 09LAI3
AC 09LAI3:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
   "Fragment C of Tetanus Toxin."
RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RL EMBL: AF154828; AAF73267.1; -.
DR HSSP; P04958; IABD.
DR InterPro; IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 451 AA: 51823 MW: 69ABC5F030E6CD8E CRC64:

```

```

Query Match 95.3%; Score 2340; DB 2; Length 451;
Best Local Similarity 98.0%; Pred. No. 3e-131;
Matches 442; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 13 KNIDCWVNDNEDIDVILKSTIILNDIINDIISDISGFNSSVITYPDAQLVPGINGKAIH 72
   |||||
Db 1 KNIDCWVNDNEDIDVILKSTIILNDIINDIISDISGFNSSVITYPDAQLVPGINGKAIH 60
QY 73 LVNSESSEVIVHKAMDEYNDEMNFVFWLVRPVASSHLEQVPTNEYSIISSMKKSHS 132
   |||||
Db 61 LVNSESSEVIVHKAMDEYNDEMNFVFWLVRPVASSHLEQVPTNEYSIISSMKKSHS 120
QY 133 LSTSGWSVSLKGNLIMTLKDSAGEVROITFRDLDPKFNAYLANKWVITITDRILSSA 192
   |||||
Db 121 LSTSGWSVSLKGNLIMTLKDSAGEVROITFRDLDPKFNAYLANKWVITITDRILSSA 180
QY 193 NLVINGVMSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFCALNPKETIEKL 252
   |||||
Db 181 NLVINGVMSAEITGAIREDNNITLKLDRCNNNOVYSIDKFRIFCALNPKETIEKL 240
QY 253 YTSLSITFLRDEFGNPLRDYETLYLIPVASSKDVOLKNTTYMYLTNAPSTYNGKLN 312

```

```

Db 241 YTSLSITFLRDEFGNPLRDYETLYLIPVASSKDVOLKNTTYMYLTNAPSTYNGKLN 300
   |||||
QY 313 YRRLVGLKFEIIRKTPPNEIDSFVKSQDFIKLYSVNNNEHIVGYPKDGNAFNNIDRI 372
   |||||
Db 301 YRRLVGLKFEIIRKTPPNEIDSFVKSQDFIKLYSVNNNEHIVGYPKDGNAFNNIDRI 360
QY 373 LRVGVNAPGIPLYKKMEAVKLRDLKITYSVOLKLYDDKNASIGLVGTINGQIGNDPNRDI 432
   |||||
Db 361 LRVGVNAPGIPLYKKMEAVKLRDLKITYSVOLKLYDDKNASIGLVGTINGQIGNDPNRDI 420
QY 433 IASNMVYNNHLKDKITLGCDDWYFVPTDEGWTND 463
   |||||
Db 421 IASNMVYNNHLKDKITLGCDDWYFVPTDEGWTND 451

```

## RESULT 3

```

008077 PRELIMINARY: PRT: 1291 AA.
ID 008077
AC 008077:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE Botulinum neurotoxin type B (EC 3.4.24.-) (BONT/B).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EKLUND 17B AFCC25765;
RX MEDLINE=94122659; PubMed=7764370;
RA Huison R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic
RT Clostridium botulinum type B neurotoxin: comparison with other
RT Clostridium neurotoxins."
RL Curr. Microbiol. 28:107-110(1994).

```

```

CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOSOMAL-2.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN. TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC NEUROTOXIN.
CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC REGION.
DR EMBL: X71343; CA450482.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTPeptidse.
DR Pfam; PF01742; Peptidase_M27.1.
DR PRINTS; PR00760; BONTTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin.1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW SEQUENCE 1291 AA: 150513 MW: 71BCAFED23D69FAA CRC64:

```

```

Query Match 25.1%; Score 617; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 1.9e-28;
Matches 149; Conservative 94; Mismatches 191; Indels 46; Gaps 14;
QY 5 TTPSPYSKNIIDCWVNDNEDIDVILK-----KSTIILNDIINDIISDISGFNSV 54
   |||||

```

```

Db 834 TITPELDS-----TYSNIEILLIKFNKYNSEILNNIILNRYRDNNDLIDSGYAKV 885
QY 55 ITPPAOLVPGINGK-AIHLVNNESSEVYIHKAMDIEYNMFNFTVSEFLRVPKVSASH 113
Db 886 EYVDGVLK-----NDKNOFKLTSSADSRIKTYQONONIFNSMFLDEFSVFWIRIKRYRND 941
QY 114 LRGYGNESYIISMKKSHLSIGSGSVSLKGNLWTLKDSAGEVQIIFR-DLPKFN 172
Db 942 IONVTHNEYTIINCKMKN-----SGWKISIRGNMIITWLLIDINGKIKSVFEFYIIRDIS 996
QY 173 AVLANKVFTITINDRLSSANLYINGVLMSAETLTGALREDNNTILKDRCNNNOYV 232
Db 997 EXI-NMFEVITINN-LDNAKITINGTLESNMDIKDIGEYIVGEITFKLDGDVDRQFI 1054
QY 233 STDKRIFICKALNPKIEIKLYSYSLITFLRDWGNPLRYDTEYLYIPVASSKDVOLKN 292
Db 1055 WKYPSIFNTQLOSNIKEIKYQSYSEYLLKDEWGNPLMKNKEYYEMNAGKNKSYIKLVK 1114
QY 293 ITPDYMVLTPASVTNGKLNYYRRLYNGLKFTIKRTYPPNET-DSEYKSGDFIKLYSVN 351
Db 1115 DSSVEIILRSKYNNONSNTYINRNLTIGERFIIIRRESNOSINDDIYRKEDYIHLDLVLH 1174
QY 352 NNE-HIVGPPKGNAPNNLDRILRVGYNAPGIPLYKMEAVKLRDLKTYSVOLKVDKKN 410
Db 1175 HHEMRYVAY-----KYFKKEEKLFLISDSNFEYTIIEIKEDQPSYCOLFFKDEE 1230
QY 411 AS--LGLVGTGTH--NGOIGNDPNRDLIASNMYFNHLKDK---ILGCDWYFVPTDEGWT 461
Db 1231 STDIGLIHMFYESGVLKRRKKYKDYFCISKWYLKEVKRKYPSKNGCMWQFIKDEGWT 1290

RESULT 4
Q933KO PRELIMINARY: PRT: 1291 AA.
AC 0933KO:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
OS Type B cryptic neurotoxin.
OC Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RN SEQUENCE FROM N.A.
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RA STRAIN=593, AND 588;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contain type B toxin gene sequences."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF300465; AAL11498.1; -.
DR EMBL; AF300465; AAL11498.1; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR Neurotoxin.
SQ SEQUENCE 1291 AA: 150843 MW: 7AC1737B0FA5A151 CRC64:

Query Match 24.4%; Score 600; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 2e-27;
Matches 148; Conservative 96; Mismatches 193; Indels 40; Gaps 15;
QY 5 TPIPSYKNDLCWDNEDIDVILK-----KSTINLDINDNDISDISGNSVITYP 58
Db 834 TSIPF-----DLSTYNTNLTILIEFNKYNSDILNNIILNLRNKLIDLSGYAKAVEYD 889
QY 59 DAOLVPGINGK-AIHLVNNESSEVYIHKAMDIEYNMFNFTVSEFLRVPKVSASHLEQY 117
Db 890 GVKL-----NDKNOFKLTSSANSKIRVYQONONIFNSMFLDEFSVFWIRIKRYNDGIQNT 945
QY 118 GTNEXYIISMKKSHLSIGSGSVSLKGNLWTLKDSAGEVQIIFR-DLPKFNAYLA 176

```

```

Db 946 IHNEYTIINCKMKN-----SGWKISIRGNMIITWLLIDINGKIKSVFEFYIIRDIS 999
QY 177 NKWVFTITINDRLSSANLYINGVLMSAETLTGALREDNNTILKDRCNNNOYVSIK 236
Db 1000 NMFEVITINNS-DNAKITINGKLESHIDIRDIERYIANDELIFKLDGNDRQFIWMKY 1058
QY 237 FRIECKALNPKIEIKLYSYSLITFLRDWGNPLRYDTEYLYIPVASSKDVOLKNITDY 296
Db 1059 FSIFNTLSOSNIEEIKYQSYSEYLLKDEWGNPLMKNKEYYEMNAGKNKSYIKLVK 1118
QY 297 MYLTNAPSVTNGKLNYYRRLYNGLKFTIKRTYPPNET-DSEYKSGDFIKLYSVYNNNE 354
Db 1119 GEILRSKYNNONSNTYINRNLTIGERFIIIRRSNOSINDDIYRKEDYIHLDEFNLQEW 1178
QY 355 HIVGPPKGNAPNNLDRILRVGYNAPGIPLYKMEAVKLRDLKTYSVOLKVDKKNAS-- 412
Db 1179 RVYMY-----KYFKKEEKLFLAPISDSDEFYNTIQIKEDQPTYSQOLLFKDESTDE 1234
QY 413 LGLVGTGTH--NGOIGNDPNRDLIASNMYFNHLKDK---ILGCDWYFVPTDEGWT 461
Db 1235 IGLIGHMFYESGVLKRRKKYKDYFCISKWYLKEVKRKYPSKNGCMWQFIKDEGWT 1290

RESULT 5
Q92AJ8 PRELIMINARY: PRT: 1291 AA.
AC 092AJ8:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Bont Protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RN SEQUENCE FROM N.A.
RA STRAIN=CDC 3281;
RC MEDLINE=98440323; PubMed=9767710;
RA Santos-Buejiga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins."
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL; Y13630; CAAT3968.1; -.
DR HSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SEQUENCE 1291 AA: 150840 MW: E4D3B0E46AB2E735 CRC64:

Query Match 24.4%; Score 599; DB 2; Length 1291;
Best Local Similarity 31.0%; Pred. No. 2.3e-27;
Matches 148; Conservative 96; Mismatches 193; Indels 40; Gaps 15;
QY 5 TPIPSYKNDLCWDNEDIDVILK-----KSTINLDINDNDISDISGNSVITYP 58
Db 834 TSIPF-----DLSTYNTNLTILIEFNKYNSDILNNIILNLRNKLIDLSGYAKAVEYD 889
QY 59 DAOLVPGINGK-AIHLVNNESSEVYIHKAMDIEYNMFNFTVSEFLRVPKVSASHLEQY 117
Db 890 GVKL-----NDKNOFKLTSSANSKIRVYQONONIFNSMFLDEFSVFWIRIKRYNDGIQNT 945
QY 118 GTNEXYIISMKKSHLSIGSGSVSLKGNLWTLKDSAGEVQIIFR-DLPKFNAYLA 176
Db 946 IHNEYTIINCKMKN-----SGWKISIRGNMIITWLLIDINGKIKSVFEFYIIRDIS 999

```

[illegible]

RESULT 6	
Q93G71	
ID Q93G71	PRELIMINARY; PRT; 1291 AA

01-DEC-2001 (TReMBLrel. 19, Created)  
 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Neurotoxin type B.  
 OS Clostridium botulinum  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1436;  
 RA Kirma N., Ferreira J.L., Baumstark B.R.;  
 RT "Characterization of six type A strains of Clostridium botulinum that  
 RT contain type B toxin gene sequences.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF295926; AAK97132.1; -;  
 DR InterPro: IPR000395; BontTox1stn.  
 DR InterPro: IPR00130; Zn\_MPEpdase.  
 DR Pfam: PF01742; Peptidase\_M47; 1.  
 DR ProDom: PD001963; BontTox1stn; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 SQ SEQUENCE 1291 AA; 150824 MW; D7CA07BAEE2B8CD2 CRC64;

24.38; Score 597; DB 2; Length 1291;

Best Local Similarity 31.08; Pred. No. 3e-27;  
Matches 148; Conservative 96; Mismatches 193; Indels 40; Gaps 15

[illegible]

```

D6 1119 GEILITRSKIYOMNSKJINYNRDLIYGEFFITIRRSNOSGINDIDYVRKEDYIYDFFNLQEW 1178
QY 355 HIVQPKDGNAPNNLDRLILRVGINAPGIPLYKKMEVNLRLDKYTSYQKLXDOKNAS-- 412
D6 1179 RVIYIY---KYFKKEEKLFLAFIYSOSDEFYNTIQKEYDEOPTYSCQLFKKDESTE 1234
QY 413 IGLVIGTH---NSOIGDMPDRDILLIASNNVFNLJDKR---ILGCDWVFPVPPDEGWT 461
D6 1235 IGLIGTHRFESGIVRE-KYDFCJASKWYLRKEVRRKPYNSKIGCMQFIPKDEGWT 1290

```

RESULT 7	
Q45851	
ID Q45851	PRELIMINARY; PRT: 1268 AA

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Neurotoxin type F.  
GN BONT / F.  
OS Clostridium baratii.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1561;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93252228; PubMed=9486245;  
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,  
RA Richardson P.T.;  
RT "Nucleotide sequence of the gene coding for Clostridium baratii type F  
RT neurotoxin: Comparison with other clostridial neurotoxins.";  
RL FEMS Microbiol. Lett. 108:175-182(1993).  
DR EMBL: X68262; CAA48329.1; -  
DR HSP; P10845; 3BTA.  
DR MEROPS: M27.002; -  
DR InterPro: IPR000395; Bontoxilysin.  
DR InterPro: IPR000130; Zn\_MTPepidase.  
DR Pfam: PF01742; Peptidase\_M27; 1.  
DR PRINTS: PR007460; BONTOXILYSIN.  
DR ProDom: PD001963; Bontoxilysin; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match	Score	DB 2;	Length
23.28;	571;	DB 2;	1268;

Best Local Similarity 32.1%; Pred. NO. 1e-25;  
Matches 153; Conservative 90; Mismatches 168; Indels 66; Gaps 19;

0Y 7 IPFSTSKNLDWCWNEEDIDVILK-----KSTILNDINDIISDISGFNSSVIYPDA 60  
||| : ||| : | ||: ||: |  
| : : : : ||| : |

Db 833 IPFELSE---YTNDKILHLIRFYKRIIDSSILMKIYENNRIFIDSSGYSNISINGDI 88

Qy 61 QLVPGINGKAITHLVNNESESVIVHKAMDIEYNMENNFTYSFMLRPRVSAASHDEQYGIN 12

889 YIY-SINRNQGIYSSRLSEINITQNNITLIYNSRIQNFVSFVWKLPAI--INLNLN IN 24

045 EYTIITMOPNN-----SGWTSI,NYNNITWTLOITGNNOKLVEFYOMIDISD---YI 99

DU	343	ELLIMCHANN	CONFIDENCE	23
DU	176	ANKWFTTNDRISSAN	YINGVMSAEITGCAIREDNNIT	KLDRCNNNOYVSID

Db 997 - NKWETVITNNRGLGSKLYINGNLTDOKSLNLCNIHVDDNLLFKIYGC - NDRVVGIR 10

QY 236 KRFIFKALNPKETIEKLYTSYLSITFLRDEFGNPIRYDTEYLI ---PVASSKDYOLK 2

```
Db      1055 YKFIFFNEIDKTEIETLYHSEPDSTILKDWFNGNYLLINKRYILLNLKPNMSVTAKSDIL 1
```

QY 292 NITDYMYLTNAPSTNGKLNLYR-RLXNGLKFITIKRY-TPNNEIDSEVKSQDFTIKLYV 3

```
Db 1115 NI-----NRQRCYSKTNIFSNARLYTGVVEVILKKGSTDTNTDNFEVRKNDTYINV 1
```

QY 349 SYNNEHIVGYPKDGNF---NDRILRVGNAPGIPLYKKNEAVKLRDLKTYSSVQLK 4

\_\_\_\_\_



```

DB 1168 YDGNSEYQYADVSTSAVEKTIKLRRISNSNYSNOMLIMDSIGDCTMNFKT----- 1220
OY 406 YDDKNAISLGLVGTHTNGQINDPNDRLILASNMWFNLHKDIL--GDMYVPPVDEGW 460
DB 1221 --NNCGNDIGLGFH---LNN-----LVASSWYKKNIRNTRNGCFWSPISKEHGW 1266

RESULT 8
AC 09K395 PRELIMINARY: PRT: 1251 AA.
ID 09K395:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Type E botulinum toxin.
GN BONT/E.
OC Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1492;
RN SEQUENCE FROM N.A.
RP STRAIN=LCL 095;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 095) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=LCL 155;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
RA Kato H., Nakamura S., Karasawa T.;
RT "C. butyricum (LCL 155) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=KZ 1899;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=KZ 1887;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=K2 1898;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1898) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=K2 1886;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1886) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=K2 1887;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1887) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=K2 1889;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1889) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1889) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1890) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=K2 1891;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (K2 1891) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN SEQUENCE FROM N.A.
RP STRAIN=LCL 063;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 063) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB037714; BAB03522.1;
DR EMBL: AB037704; BAB03512.1;
DR EMBL: AB037705; BAB03513.1;
DR EMBL: AB037706; BAB03514.1;
DR EMBL: AB037707; BAB03515.1;
DR EMBL: AB037708; BAB03516.1;
DR EMBL: AB037709; BAB03517.1;
DR EMBL: AB037710; BAB03518.1;
DR EMBL: AB037711; BAB03519.1;
DR EMBL: AB037712; BAB03520.1;
DR EMBL: AB037713; BAB03521.1;
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN; 1.
SQ SEQUENCE 1251 AA; 143751 MW; 2021F4E427070296 CRC64;

Query Match 23.28; Score 570.5; DB 2; Length 1251;
Best Local Similarity 30.08; Pred. No. 1,le-25;
Matches 150; Conservative 84; Mismatches 161; Indels 105; Gaps 17;

OY 1 MFEST---PIPFYSKMLDCWVNDDEDIVL-----KKSTLNLNDINDISD 46
DB 814 MVIDTLNNSPTFKLSSTYD-----DKLILSYFNKFFRIKSSVLMRYKNDKYVD 864
OY 47 ISGFNSVITYTPPAQLVPGINGKAIHLVNNSESEVIVKAMDIEYNFNFTVSFWLRY 106
DB 865 TSGYDSNNININGEIFIYP-TNKNQFTIENSKPSEVNISQNDYIIVDNKKNSISFWVAI 923
OY 107 PKYSASHLEQGTNEYISLISMKKHSLSGSGWSYSLKGNLIMPLKLSAGVQITFR- 165
DB 924 PNYDNKIVNT--NNEYITLNCMRDNN---SGWKVSLNHNELIWTLODNARNOKLVKY 977
OY 166 DLDPKFNAVYLANKWVFITITNDRSLSSANLYINGVLMGSAETTGAIREDNNITLKDRC 225
DB 978 GNNNGISDYI-NKWIETITNDRIGDSKLYINGHLIDOKSLNLGNITVSDITLKIYNC 1036
OY 226 NNNQVYSIDKFRICFALNPKETLEKLYSTLTFLEDFWGNPLRYDTEYVLLIPVASS 285
DB 1037 -SYTRYIGIRYFNFDKELDETETIQTLYSNPNINILKDFGNYLLYDKGYLLAVLPN 1095
OY 286 -----KD--VOLKNITDYMVLTNAPSTYTNKLNIIYRLKNGLKFTIKRTYNNDEISF 337
DB 1096 NFDRRKDSSTLSTNNISSTILLAN-----RLYSGIKVKIQRVNDSSSTNDRF 1141

```

QY	338	VKSGCFILVYSYNNN-----EHVYGRKDGNAFNNDRLRLVGYNAAGI	382
Db	1142	VRKND--GVYINYSNSSSYLYADPTNTTDEKTKTSSSSGRFRQVYVMSYVNNC---	1196
QY	383	PLYKMEAVKLRDLTKYVSQVLTKYDOKNASIGLVYTHNGQIGNDPNRDILASNNYFNHL	442
Db	1197	-----TMMFRKNNNNNICLLGF-----KADTVVASTWYTYTHM	1228
QY	443	KD-KILGCDWYFVPTDEGM	460
Db	1229	RDHTNSGCFWNFISEHGM	1248
RESULT 9			
ID	Q9X708	PRELIMINARY;	PRT; 441 AA.
AC	Q9X708;		
DT	01-NOV-1999 (TREMBlrel. 12, Created)		
DT	01-NOV-1999 (TREMBlrel. 12, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	Botulinum neurotoxin type B (fragment).		
BO	BOYT/B.		
OS	Clostridium botulinum.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;		
OC	Clostridiales; Clostridiaceae; Clostridium.		
OX	NCBI_TaxID=1491;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99343691; PubMed=10413679;		
RA	Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,		
RA	Schiavo G.;		
RT	"Functional characterisation of tetanus and botulinum neurotoxins		
RT	binding domains."		
RL	J. Cell Sci. 112:2715-2724(1999).		
DR	EMBL; AJ242628; CAB43706.1; "		
DR	HSSP; P10845; 3BTA.		
KM	Neurotoxin.		
FT	NON_TER	1	441
FT	NON_TER	441	441
SO	SEQUENCE	441 AA; 52772 MW; 721D0B46BE8C95A4 CRC64;	
Query Match			
	Best Local Similarity	23.1%; Score 567; DB 2; Length 441;	
	Matches 138; Conservative	89; Mismatches 179; Indels 28; Gaps 12	
QY	34	ILNLDINDIISDISGENSSVYTPDAQLVPGINGK-AIHLVNNSSSEVYVHKADIEYN	92
Db	23	INLRYRDNNLIDLISGAGVVEYDDEL---NDKNGCFKISTSTSETRVQNONIIFN	78
QY	93	DMEFNNTVSEFWLRVPRVSAHSLEQVYNTNYSIISSMKKSLSTGSGWSYSLGNNLIWTL	152
Db	79	SMEFLDSEVSEFWLRIPKYNKDGIIONYTHNETYIINCINN---SGWKISIKGNRIWTL	133
QY	153	KSGASGVROITER-DLPDFENAVIANKWVFITITNDRLSSANLIYNGVLSAETIGLA	211
Db	134	TDINGTKSVFEEYSIREDISDI-NRMFVITINNS-DNAKIYINGKLESINDIDICE	191
QY	212	IKEDNNITLKLDRCCNNNOYVSIDKRFICKALNPEIEKLTYSITFLDFQGNPLR	271
Db	192	VIANGEIIFKLDGDDIDRTOFIWAKYFSINTELOSNIKEIYKIOSYSELDKDFWNPDM	251
QY	272	YDVEEYLLIPVASSKVOVLKNTIDWYILNAPSYTGKLNIIYRRLYNGKLFIRKRYTPN	331
Db	252	YNEEYMFNAGNNSYIKLAKDSSVGEILTRSKYINONSNYINRNLITGEKFIIRKSSNS	311
QY	332	NEI-DSFVSGDFIKIYVSYNNNEHIVGPKDGNAFNNNDRLRVGYNAPG.IPLYKMEA	390
Db	312	QSLNDIVRKREDYIYIDFNSNRNEMRYAVKD--EKEEKKLVLANIYDSNEFYKTIOI	368
QY	391	VKLRDLKTVSVOLKLYDOKNAS--LGLVGH----NGQIGNDPNRILIASNNYFNHL	444
Db	369	KVDEDPYSCQLLFKKDESTEIDGIGIHRYESGIVLRD--YKNVFCISKRYLAEVVR	427

OY	445	K--ILGDMWFEV 454           :           :
Db	428	KPYNPNGCMMOFT 441
RESULT 10		
ID	OPFAR6	PRELIMINARY; PRT: 1255 AA.
AC	OPFAR6:	
Dt	01-MAR-2001 (TREMBLrel. 16,	Created)
Dt	01-MAR-2001 (TREMBLrel. 16,	Last sequence update)
Dt	01-DEC-2001 (TREMBLrel. 19,	Last annotation update)
DE	Type E botulinum toxin.	
CN	BONT/E	
OS	Clostridium butyricum.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;	
CC	Clostridiales; Clostridiaceae; Clostridium.	
OX	NCBI_TaxID=1492;	
RN	[ ]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN=BL_6340/ATCC 43755/BL_5520/KZ_147;	
RX	MEDLINE=20509829; PubMed=11059594;	
RA	Yang X., Maegawa T., Kurasawa T., Kozaki S., Tsukamoto K., Gyoju Y.,	
RA	Wamkawa K., Oguna K., Sakaguchi Y., Nakamura S.;	
RT	"Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium	
RT	butyricum Strains."	
RL	Appl. Environ. Microbiol. 66:4992-4997(2000).	
DR	EMBL: AB039264; BABI2249.1.-.	
DR	HSSP: P10845; 3BTA.	
DR	Interpro: IPR000395; Bontoxilysin.	
DR	Interpro: IPR000130; Zn_Mpeptase.	
DR	Pfam: PF01742; Peptidase M27; 1.	
DR	PRINTS: PR00760; BONTMOTILYSIN.	
DR	Prodom: PD001963; Bontoxilysin; 1.	
DR	PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.	
SQ	SEQUENCE 1255 AA; 143918 MW; 1B57B9D85CD8E4D CRC64;	
Query Match 23.0%; Score 565; DB 2; Length 1255;		
Best Local Similarity 29.3%; Pred. No. 2,3e-25;		
Matches 147; Conservative 82; Mismatches 166; Indels 106; Gaps 16.		

```

OY 442 LKDKI--LGCWYFVPTDEGR 460
Db 1232 MRDNTNSNGFFWFISEHGW 1252

RESULT 11
O57236 PRELIMINARY; PRT; 1278 AA.
ID 057236
AC 057236; Q45863;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Botulinum neurotoxin type F (BONT/F protein).
BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1491;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 10281;
RA Hutson R.A., Collins M.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
RN 12)
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 10281;
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN 13)
RP SEQUENCE OF 635-1000 FROM N.A.
RC STRAIN=NCCT 1028;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN 14)
RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=LANGEAND;
RX MEDLINE=98404102; PubMed=9732534;
RA East A.K., Bhandari M., Hielm S., Collins M.D.;
RT "Analysis of the botulin neurotoxin type F gene clusters in
RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT batelli."
RL Curr. Microbiol. 37:262-268(1998).
DR EMBL; X81714; CAA57358.1; -
DR EMBL; L35496; AAA23210.1; -
DR EMBL; X70821; CAA50152.1; -
DR EMBL; X99064; CAA67512.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR ProSite; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 22.5%; Score 553; DB 2; Length 1278;
Best Local Similarity 30.0%; Pred. No. 1,2e-24;
Matches 138; Conservative 92; Mismatches 152; Indels 78; Gaps 18

OY 29 LKSTILINDNNIIFISGFSNSVITPPDAQLVPGINGKAIHLVNNSESEYIYHKAMD 88
Db 866 IKDMSIIDMRKNNKFDISGYSNISTINGVYIT-STNRNQFGIYSSKSPSEVNAQNND 924

OY 89 IEYDNMFNETVSEFWLRYPKVYSASHLEQYG-TNEYSIISMKKHSLSIGSGSVSLKGN 147
Db 925 IYYGRYQNSISFWRPK---YFKVNLNNEETIIDLCRRNN---SGKYSILNTNK 976

OY 148 LIWILKKSAGSVRIITR-----DLPRKFNAYLANKWVEYITITNDRISSANLYINGVLNG 202

```

```

Db      977  IETWLTQDTAGNKKQALFEVYVYQTMISTSD-----YT-NKMLFVATTNRLNLSRYITNGNLTID 1031
QY      203  SAEITTGCAIFREDNNITILKIDRCNNNNQVSTIDKRFIPCKALNPKNEIKLYTSYSTIFPL 262
Db      1032  EKSISNIGDIIHVSNIIFKIVGC-NDFRYVGIRFKFYDTELGTKEIETLVSDDEPSPIL 1090
QY      263  RDPGNPLPRDYETELLYLPVASSKDV-QLKTIITDYMILTAPSTTNGKLNLIYR-RLYNG 320
Db      1091  KDEFGNELLINKRYILLNLFTDSTSTQNSMFLN-----INQORGYOKQPNIFPSTRLYTG 1146
QY      321  LKFLIKR--YTPNNEIDSFVKSDF-----IKLY--VSYNNEHIVGVPKQDA 365
Db      1147  VEVIIRKNGSTDSINTQNPFRKNDLAIYNDVDREYRLYADISIAPEKIIKILIRFNS 1206
QY      366  FNNIDRL---RQYNAPIGLYRKMEAVKLRDLKTYSVOLKLYDDKNASLGVLGTHNQ 422
Db      1207  NNSIGQIIVMDSITGNC-----TWNFQNNNGNIGLGLGFSNN 1244
QY      423  IGNDPNEDILIASNMWYFNHLKDKIL--GCDWYFVPTDEGW 460
Db      1245  -----LVASSWYNNIRKNTSSNGCFWFSISKDEHW 1275

RESULT 12
Q9ZAJ5 PRELIMINARY: PRT; 1280 AA.
ID Q9ZAJ5
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Bont protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
RN NCBI_TaxId=1491;
RX [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CDC 3281;
RA MEDLINE=98440323; PubMed=9767710;
RA Santos-Buelga J., COLLINS M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of Clostridium botulinum producing type B & F
RT neurotoxins."
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13631; CAA73972.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MPeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SO SEQUENCE 1280 AA; 147487 MW; D0F748976EBC222C CRC64;

Query Match 21.7%; Score 532.5; DB 2; Length 1280;
Best Local Similarity 29.2%; Pred. No. 2e-23;
Matches 140; Conservative 89; Mismatches 137; Indels 113; Gaps 20;

QY 29 LKSTITLMDLNNDIISIGSGNSSVITYPDAQIVPGINGKA-IHLVNNES-----S 79
Db 866 IKDSIIILDMRENNKFIIDISGYSNI-----SINGNYIYSTNRNQGIISSGRIS 915
QY 80 EVIYHAKADIYNNMFNFTYSFMLRVPKYSASLHEQYGTN-EYSIISMKKHSLSISG 138
Db 916 EVNIAQNNDIITYNSKYNFSISFWTIPK---HYRPNRRNREYTIINCMMNNN---SG 967
QY 139 WSVSL---KGNLNIWTLKDSAGEVQIIFR-DLPDKFAAYLANKKVFETITDRLSANTL 194
Db 968 WKISLRTDCEIITWLTQDTSGNKKELFRKEBELASISDYI-NKRIEYVTTIYNNRIGNSRI 1026

```



```

Db 1059 LMRDNIKSKELSNEDINIVEGQILRNVIKDYWGPNLKFDEYYII--NDNYIDRYIA 1116
QY 292 NITDYMVLTPNAPSYTNGKLNITYRRLYNGLKFLIKRTPNNEIDSEVKSQDFTKLYSYN 351
Db 1117 PKSNITLVLYOYFDRS-----KLYTGNPTITIKSVSDKNPY--SRILNGDNIMFHLN 1166
QY 352 NNEHIVGYPRKDGNAFNNLDRIILRVGNAPGIPLYKK-MEAVKLR-DLKTYSVOLKLYDDK 409
Db 1167 SKGYMI-----IRDDTPI-----YALGREGCKNCVYALKQSNLNGYGI--GIFSIK 1212
QY 410 NASLGLVGTN--GOIGNDPNRDIILASN--WYFN-----HLKDKILGCD-- 450
Db 1213 N----IVSQNKYCSQIFSSFMKNTMLLADIKPWRFSFENAYTPVAVTNYETKLLSTSSF 1268
QY 451 WYFVPTDEGW 460
Db 1269 WKFISSRDPGW 1278

```

RESULT 15

```

Q45849 ID 045849 PRELIMINARY: PRT: 1280 AA.
AC 045849:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Neurotoxin.
OS Clostridium botulinum C.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_taxid=36628;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6813;
RX MEDLINE=96156810; PubMed=8593068;
RA Moritski K., Koura M., Fujii N., Fujinaga Y., Inoue K., Syuto B.,
Oguma K.;
RT "Molecular cloning of the gene encoding the mosaic neurotoxin,
RT composed of parts of botulinum neurotoxin types C1 and D, and PCR
RT detection of this gene from Clostridium botulinum type C organisms.";
RL Appl. Environ. Microbiol. 62:662-667(1996).
DR EMBL: D49440; BAA08418.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Metpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR ProSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SO SEQUENCE 1280 AA; 147815 MW; 74F813B28B8C989 CRC64;

```

Query Match 17.0%; Score 417.5; DB 2; Length 1280;  
 Best Local Similarity 26.1%; Pred. No. 1.3e-16;  
 Matches 128; Conservative 100; Mismatches 179; Indels 83; Gaps 20;

```

QY 3 FSTPIPFYSKMLDCWVNEEDIVY-----LKKSTIILNDIINDIISDGFNSVIT 56
Db 840 FENLIPF-----NIFSYTNLSLKDIINEFNSINDSKILSLQKKNAALVDTSGYNAEVL 895
QY 57 YPDADLVGKNGKAIHL-----VNNESSEVIYKRAMDIEYNDMFNFTVSFWLVRPK-VSA 111
Db 896 EDDVYV-----NTIYTNDFKLSSSGDKIIVLNNNILEYSAIYENSVSFWIKISKDLTN 949
QY 112 SHLEQYGTNEYSIISMKHSLSTGSGWSYSLKGNLILWTLKDSAGEVQIITFRDLPRK 171
Db 950 SH-----NETIISIKON-----SGWKLCLIRNGNIEIWLIDINRKYSLLIFDYSESLS 998
QY 172 NAYLANKWVFTITINDRLSANYLYNGVLMGSAEITGLGAIREDNNITLKLDRCNNNOY 231
Db 999 HTGYTNKKWFYITITNINMGYMKLYINGELKQSERIEDLNEVAKLDKTIIVGIDENIDENOM 1058

```

```

QY 232 VSIDKRIEPCALNPKETIEKLYTSYLSITFLRDFWGNPLRYDTEYLLIPVASSKDVOLK 291
Db 1059 LMRDNIKSKELSNEDINIVEGQILRNVIKDYWGPNLKFDEYYII--NDNYIDRYIA 1116
QY 292 NITDYMVLTPNAPSYTNGKLNITYRRLYNGLKFLIKRTPNNEIDSEVKSQDFTKLYSYN 351
Db 1117 PKSNITLVLYOYFDRS-----KLYTGNPTITIKSVSDKNPY--SRILNGDNIMFHLN 1166
QY 352 NNEHIVGYPRKDGNAFNNLDRIILRVGNAPGIPLYKK-MEAVKLR-DLKTYSVOLKLYDDK 409
Db 1167 SKGYMI-----IRDDTPI-----YALGREGCKNCVYALKQSNLNGYGI--GIFSIK 1212
QY 410 NASLGLVGTN--GOIGNDPNRDIILASN--WYFN-----HLKDKILGCD-- 450
Db 1213 N----IVSQNKYCSQIFSSFMKNTMLLADIKPWRFSFENAYTPVAVTNYETKLLSTSSF 1268
QY 451 WYFVPTDEGW 460
Db 1269 WKFISSRDPGW 1278

```

Search completed: December 1, 2002, 11:23:47  
 Job time : 90 secs

